



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 179670**

**TO: Satayanarayana Gudibande**  
**Location: REM/3C04/3C18**  
**Art Unit: 1654**  
**Wednesday, February 15, 2006**  
**Case Serial Number: 10/602394**

**From: Barb O'Bryen**  
**Location: Biotech-Chem Library**  
**Remsen 1a69**  
**Phone: 571-272-2518**

**barbara.obryen@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:01:37 ; Search time 116 Seconds  
(without alignments)  
37.877 Million cell updates/sec

Title: WRF

Perfect score: 58

Sequence:  ydwzfnafxy-103

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*
- 9: geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	98.3	10	9	Adw04630 Human AGR
2	57	98.3	10	9	Adw04629 Human AGR
3	46	79.3	2133	8	Adp25432 Plasmodiu
4	43	74.1	484	8	Adx80139 Plant ful
5	42	72.4	10	9	Adw04632 Human AGR
6	42	72.4	10	9	Adw04631 Human AGR
7	40	69.0	9	8	Adp25709 Plasmodiu
8	40	69.0	1804	8	Adh56613 Deduced p
9	40	69.0	1856	2	Aay21801 B. subtil
10	40	69.0	1856	3	Aay83269 Polypepti
11	39	67.2	666	4	Abb58019 Drosophil
12	39	67.2	1006	4	Abb70749 Drosophil
13	39	67.2	1364	3	Aab42843 Human ORF
14	39	67.2	1365	9	Ady19343 PRO polyp
15	39	67.2	1368	8	Adl82927 Human PRO
16	39	67.2	1368	9	Ady18160 PRO polyp
17	39	67.2	1368	9	Ady17490 PRO polyp
18	39	67.2	1368	9	Aeb86622 Human med
19	39	67.2	1439	4	Abb60094 Drosophil
20	38	65.5	315	7	Abb602364 Klebsiell
21	38	65.5	400	8	Adh52878 Bovine ap
22	37	63.8	135	8	Adk46463 Streptoco
23	37	63.8	152	8	Adr96377 Novel S.
24	37	63.8	152	9	Aea60247 Streptoco

25	37	63.8	233	6	ABU02359	Abu02359 S. pneumo
26	37	63.8	409	8	ADN26924	Adn26924 Bacterial
27	37	63.8	440	6	ABU20398	Abu20398 Protein e
28	37	63.8	618	9	ABM91016	Abm91016 M. xanthu
29	37	63.8	643	8	ADS27839	AdS27839 Bacterial
30	36.5	62.9	763	9	ABE48788	AbE48788 Pichia pa
31	36	62.1	184	8	ADX69103	Adx69103 Plant ful
32	36	62.1	191	6	ABM69582	Abm69582 Photorhab
33	36	62.1	212	8	ADx67448	Adx67448 Plant ful
34	36	62.1	276	8	ADV89584	Adv89584 Streptoco
35	36	62.1	276	8	ADV82991	Adv82991 Streptoco
36	36	62.1	276	8	ADV80837	Adv80837 Streptoco
37	36	62.1	498	7	ABO65670	AbO65670 Klebsiell
38	36	62.1	499	8	ADN17381	Adn17381 Bacterial
39	36	62.1	543	4	ABB63767	Abb63767 Drosophil
40	36	62.1	566	2	AAR13228	Aar13228 Endogluca
41	36	62.1	1062	8	ADP25429	Adp25429 Plasmodiu
42	35	60.3	9	7	ADM32693	Adm32693 HLA bindi
43	35	60.3	9	8	ADP80034	Adp80034 Human HLA
44	35	60.3	10	9	ADW04628	Adw04628 Human AGR
45	35	60.3	17	9	ADW04636	Adw04636 Human AGR

#### ALIGNMENTS

RESULT 1  
ID ADW04630 standard; peptide; 10 AA.  
XX  
AC ADW04630;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Human AGRP/MCR agonist chimeric cyclic peptide #4.  
XX  
KW protein engineering; melanocortin receptor; AGRP; agouti related protein;  
KW obesity; gene therapy; anorectic; cyclic; melanocortin agonist.  
XX  
OS Homo sapiens.  
OS Chimeric.  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at  
FT position 9"  
FT Misc-difference 5  
FT Misc-difference 9 /note= "D-form residue"  
FT Misc-difference 9 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a  
lactam bridge with Asp at position 2"

US2004260063-A1.

23-DEC-2004.

23-JUN-2003; 2003US-00602394.

23-JUN-2003; 2003US-00602394.

(HASK/) HASKELL-LUEVANO C.

Haskell-Luevano C;

WPI; 2005-030733/03.

New chimeric peptides and templates based upon melanocortin agonist  
peptides and agouti related protein antagonist peptide, useful for  
treating or preventing conditions modulated by melanocortin receptors,  
such as obesity.

Claim 2; SEQ ID NO 5; 15pp; English.

XX CC The present invention relates to a chimeric peptide that is biologically  
 CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.

XX SQ Sequence 10 AA;

Query Match 98.3%; Score 57; DB 9; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0044;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDWRFNAPXY 10  
 Db 1 YDWRFNAPXY 10  
 |||||

RESULT 2  
 ADM04629  
 ID ADM04629 standard; peptide; 10 AA.  
 XX AC ADM04629;  
 XX DT 10-MAR-2005 (first entry)  
 XX DE Human AGRP/MCR agonist chimeric cyclic peptide #3.  
 XX KW protein engineering; melanocortin receptor; AGRP; agouti related protein;  
 KW obesity; gene therapy; anorectic; cyclic; melanocortin agonist.  
 XX OS Homo sapiens.  
 OS Chimeric.  
 OS Unidentified.

XX FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at  
 FT position 9"  
 FT Misc-difference 9 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a  
 FT lactam bridge with Asp at position 2"

XX PN US2004260063-A1.  
 XX PD 23-DEC-2004.  
 XX PF 23-JUN-2003; 2003US-00602394.  
 XX PR 23-JUN-2003; 2003US-00602394.  
 XX PA (HASK/) HASKELL-LUEVANO C.  
 XX PI Haskell-Luevano C;  
 XX DR WPI; 2005-030733/03.  
 XX PT New chimeric peptides and templates based upon melanocortin agonist  
 FT peptides and agouti related protein antagonist peptide, useful for  
 FT treating or preventing conditions modulated by melanocortin receptors,  
 FT such as obesity.

XX PS Claim 2; SEQ ID NO 4; 15pp; English.  
 XX CC The present invention relates to a chimeric peptide that is biologically  
 CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such

XX CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.

XX SQ Sequence 10 AA;

Query Match 98.3%; Score 57; DB 9; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0044;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDWRFNAPXY 10  
 Db 1 YDWRFNAPXY 10  
 |||||

RESULT 3  
 ADP25432  
 ID ADP25432 standard; protein; 2133 AA.  
 XX AC ADP25432;  
 XX DT 09-SEP-2004 (first entry)  
 XX DE Plasmodium falciparum antigen amino acid sequence SEQ ID NO:9.  
 XX KW Plasmodium falciparum; malaria parasite; antigen; malaria; immunogenic;  
 KW immune response; cytostatic; anti-HIV; virucide; hepatotropic;  
 KW antibacterial; vaccine; cancer; infectious disease; AIDS; hepatitis;  
 KW bacterial infection.  
 XX OS Plasmodium falciparum.  
 XX PN WO2004053086-A2.  
 XX PD 24-JUN-2004.  
 XX PF 08-DEC-2003; 2003WO-US038966.  
 XX PR 06-DEC-2002; 2002US-0431494P.  
 XX PA (EPIM-) EPIMMUNE INC.  
 XX PA (USNA ) US SEC OF NAVY.  
 XX PI Sette A, Doolan DL, Carucci DJ, Sidney J, Southwood S;  
 XX DR WPI; 2004-468856/44.  
 XX PT New isolated and/or purified Plasmodium falciparum polynucleotide  
 FT sequences, useful in inducing an immune response for preventing and/or  
 FT treating cancer and infectious diseases, such as AIDS, hepatitis, and  
 FT bacterial infections.

XX PS Claim 22; SEQ ID NO 9; 253pp; English.  
 XX CC The present invention describes an isolated and/or purified Plasmodium  
 CC falciparum (malaria parasite) antigen polynucleotide sequence, encoding  
 CC an immunogenic peptide. Also described: (1) a primer or detection probe  
 CC for hybridisation with a target sequence or the amplicon generated from a  
 CC target sequence comprising a sequence of at least 8-30, 35, 40, 45, 50,  
 CC 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 consecutive nucleotides of any  
 CC of the polynucleotide sequences as described above; (2) a DNA chip  
 CC comprising any of the polynucleotide sequences described above; (3) a  
 CC vector comprising a promoter operably linked to any of the nucleic acid  
 CC sequences described above; (4) a host cell transformed by the vector of  
 CC (3) or the polynucleotide described above; (5) a composition comprising a  
 CC carrier and the polynucleotide described above; (6) a method of inducing  
 CC an immune response in an individual comprising the administration of the  
 CC composition of (5) to induce an immune response; (7) an isolated  
 CC polypeptide comprising any of the amino acid sequences as encoded by the  
 CC polynucleotide described above; (8) a composition comprising a carrier  
 CC and the polypeptide of (7); (9) a method of detecting P. falciparum in  
 CC biological samples, comprising contacting a biological sample with the  
 CC isolated polynucleotide and detecting the hybridisation of the isolated  
 CC polynucleotides with nucleic acids contained in the sample; (10) a method

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:14:30 ; Search time 30.6667 Seconds  
(without alignments)  
26.959 Million cell updates/sec

Title: WRF  
Perfect score: 58  
Sequence: 1'YDWRFNAPF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/PCITUS\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	67.2	220	2	US-09-270-767-45997 Sequence 45997, A
2	38	65.5	98	2	US-09-270-767-36635 Sequence 36635, A
3	38	65.5	98	2	US-09-270-767-51852 Sequence 51852, A
4	38	65.5	315	2	US-09-489-039A-8881 Sequence 8881, Ap
5	37	63.8	135	2	US-09-583-110-2978 Sequence 2978, Ap
6	37	63.8	152	2	US-09-107-433-5012 Sequence 5012, Ap
7	37	63.8	618	2	US-09-902-540-10215 Sequence 10215, A
8	36	62.1	498	2	US-09-489-039A-12187 Sequence 12187, A
9	36	62.1	566	1	US-07-862-588B-4 Sequence 4, Appl
10	35	60.3	277	2	US-09-248-796A-15054 Sequence 15054, A
11	35	60.3	280	2	US-09-252-991A-16878 Sequence 16878, A
12	35	60.3	291	2	US-09-248-796A-16362 Sequence 16362, A
13	35	60.3	334	2	US-09-248-796A-16366 Sequence 16366, A
14	35	60.3	681	2	US-09-248-796A-16367 Sequence 16367, A
15	35	60.3	769	2	US-09-248-796A-16368 Sequence 16368, A
16	35	60.3	1476	2	US-09-817-514A-4 Sequence 4, Appl
17	35	60.3	1481	2	US-09-251-645-14 Sequence 14, Appl
18	34	58.6	62	2	US-09-134-001C-4031 Sequence 4031, Ap
19	34	58.6	144	2	US-09-155-036-12 Sequence 12, Appl
20	34	58.6	144	2	US-09-155-036-13 Sequence 13, Appl
21	34	58.6	144	2	US-09-155-036-14 Sequence 14, Appl
22	34	58.6	144	2	US-09-155-036-16 Sequence 16, Appl
23	34	58.6	144	2	US-09-866-307-12 Sequence 12, Appl
24	34	58.6	144	2	US-09-866-307-13 Sequence 13, Appl
25	34	58.6	144	2	US-09-866-307-14 Sequence 14, Appl
26	34	58.6	144	2	US-09-866-307-16 Sequence 16, Appl
27	34	58.6	145	2	US-09-155-036-15 Sequence 15, Appl

28	34	58.6	145	2	US-09-866-307-15 Sequence 15, Appl
29	34	58.6	289	2	US-09-540-236-2804 Sequence 2804, Ap
30	34	58.6	294	2	US-09-489-039A-13295 Sequence 13295, A
31	34	58.6	333	2	US-09-328-352-7700 Sequence 7700, Ap
32	34	58.6	347	2	US-08-405-271A-24 Sequence 24, Appl
33	34	58.6	380	1	US-08-153-848-40 Sequence 40, Appl
34	34	58.6	380	2	US-09-299-843A-40 Sequence 40, Appl
35	34	58.6	380	2	US-09-088-337B-40 Sequence 40, Appl
36	34	58.6	380	2	US-09-149-045-2 Sequence 2, Appl
37	34	58.6	380	2	US-09-255-518C-3 Sequence 3, Appl
38	34	58.6	380	2	US-09-170-496D-62 Sequence 62, Appl
39	34	58.6	380	2	US-09-170-496D-198 Sequence 198, App
40	34	58.6	380	4	PCT-US93-11153-40 Sequence 40, Appl
41	34	58.6	455	1	US-08-679-635A-5 Sequence 5, Appl
42	34	58.6	455	2	US-09-419-163-5 Sequence 5, Appl
43	34	58.6	456	1	US-08-679-635A-6 Sequence 6, Appl
44	34	58.6	456	2	US-09-419-163-6 Sequence 6, Appl
45	34	58.6	473	2	US-09-155-036-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-270-767-45997  
; Sequence 45997, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45997  
; LENGTH: 220  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-45997

Query Match 67.2%; Score 39; DB 2; Length 220;  
Best Local Similarity 62.5%; Pred. No. 29;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 YDWRFNAP 8  
Db 99 HDWKFNVF 106  
:|:|:|:|

RESULT 2  
US-09-270-767-36635  
; Sequence 36635, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36635  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-36635

Query Match 65.5%; Score 38; DB 2; Length 98;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 DWRNAP 8
      ||| |||
Db      3 DWRLNAP 9

RESULT 3
US-09-270-767-51852
; Sequence 51852, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51852
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51852

Query Match      65.5%; Score 38; DB 2; Length 98;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 DWRNAP 8
      ||| |||
Db      3 DWRLNAP 9

RESULT 4
US-09-489-039A-8881
; Sequence 8881, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8881
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8881

Query Match      65.5%; Score 38; DB 2; Length 315;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YDWRNFA 7
      ||| |||
Db      232 YHWRNFA 238

RESULT 5
US-09-583-110-2978
; Sequence 2978, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583.110
; CURRENT FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: US 09/107.433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085.131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051.553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2978
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2978

Query Match      63.8%; Score 37; DB 2; Length 135;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 YDWRNAP 10
      ||| |
Db      64 YDWRNAP 73

RESULT 6
US-09-107-433-5012
; Sequence 5012, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107.433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5012:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...152
; SEQUENCE DESCRIPTION: SEQ ID NO: 5012:
US-09-107-433-5012
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:15:26 ; Search time 127 seconds  
(without alignments)  
32.900 Million cell updates/sec

Title: WRF

Perfect score: 58

Sequence:  $\epsilon_i$  ydwrfnafxy 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	74.1	484	US-10-425-114-49505	Sequence 49505, A
2	43	74.1	728	US-10-424-599-189465	Sequence 189465, A
3	42	72.4	89	US-10-425-115-197701	Sequence 197701, A
4	41	70.7	1114	US-10-437-963-178103	Sequence 178103, A
5	40	69.0	1769	US-10-361-522-3	Sequence 3, Appli
6	39	67.2	84	US-10-425-115-295441	Sequence 295441, A
7	39	67.2	666	US-11-097-143-849	Sequence 849, App
8	39	67.2	1006	US-11-097-143-39039	Sequence 39039, A
9	39	67.2	1368	US-10-405-387-2	Sequence 2, Appli
10	39	67.2	1439	US-11-097-143-7074	Sequence 7074, Ap
11	38	65.5	126	US-10-437-963-200586	Sequence 200586, A
12	38	65.5	127	US-10-437-963-104836	Sequence 104836, A
13	38	65.5	287	US-10-437-963-132229	Sequence 163229, A
14	38	65.5	459	US-10-437-963-118885	Sequence 118885, A
15	38	65.5	550	US-10-437-963-161824	Sequence 161824, A
16	38	65.5	599	US-10-437-963-126846	Sequence 126846, A
17	38	65.5	604	US-10-437-963-150085	Sequence 150085, A
18	38	65.5	788	US-10-437-963-204446	Sequence 204446, A
19	38	65.5	1133	US-10-437-963-143474	Sequence 143474, A
20	38	65.5	1309	US-10-437-963-115740	Sequence 115740, A
21	38	65.5	1426	US-10-437-963-169077	Sequence 169077, A
22	38	65.5	1452	US-10-437-963-128047	Sequence 128047, A
23	38	65.5	1493	US-10-437-963-170048	Sequence 170048, A
24	38	65.5	1529	US-10-437-963-204169	Sequence 204169, A
25	38	65.5	1569	US-10-437-963-117837	Sequence 117837, A
26	38	65.5	1685	US-10-437-963-171019	Sequence 171019, A
27	38	65.5	1698	US-10-437-963-192546	Sequence 192546, A

28 38 65.5 1770 4 US-10-437-963-117838 Sequence 117838, A

29 37 63.8 49 4 US-10-437-963-184166 Sequence 184166, A

30 37 63.8 152 5 US-10-617-320-5012 Sequence 5012, Ap

31 37 63.8 233 5 US-10-472-928-3874 Sequence 3874, Ap

32 37 63.8 361 4 US-10-437-963-199806 Sequence 199806, A

33 37 63.8 409 4 US-10-369-493-9577 Sequence 9577, Ap

34 37 63.8 440 4 US-10-282-122A-48322 Sequence 48322, A

35 37 63.8 643 4 US-10-369-493-16872 Sequence 16872, A

36 37 63.8 647 4 US-10-437-963-126808 Sequence 126808, A

37 37 63.8 742 4 US-10-424-599-227627 Sequence 227627, A

38 37 63.8 1012 4 US-10-437-963-107191 Sequence 107191, A

39 37 63.8 1588 4 US-10-437-963-196997 Sequence 196997, A

40 37 63.8 1663 4 US-10-437-963-143452 Sequence 143452, A

41 36 62.1 75 4 US-10-425-115-278415 Sequence 278415, A

42 36 62.1 184 4 US-10-425-114-39946 Sequence 39946, A

43 36 62.1 212 4 US-10-425-114-39291 Sequence 38291, A

44 36 62.1 308 4 US-10-424-599-283799 Sequence 283799, A

45 36 62.1 499 4 US-10-369-493-34 Sequence 34, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-425-114-49505  
; Sequence 49505, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53323)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 49505  
; LENGTH: 484  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3051-077-E3\_FLI.pep  
US-10-425-114-49505

Query Match 74.1%; Score 43; DB 4; Length 484;  
Best Local Similarity 80.0%; Pred. NO. 91;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YDWRFNAPXY 10

DB 197 YDKRFNAPHY 206

##### RESULT 2

US-10-424-599-189465  
; Sequence 189465, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 189465

```
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142102C.1.pep
US-10-424-599-189465

Query Match      74.1%; Score 43; DB 4; Length 728;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDWRFNAPFY 10
   |||||||
Db 441 YDKRFNAFY 450

RESULT 3
US-10-425-115-197701
; Sequence 197701, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 197701
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_111887C.1.pep
US-10-425-115-197701

Query Match      72.4%; Score 42; DB 4; Length 89;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YDWRFNAPFY 10
   |||||
Db 47 YDWRMNVIAF 56

RESULT 4
US-10-437-963-178103
; Sequence 178103, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 178103
; LENGTH: 1114
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (1)..(1114)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75692C.1.pep
US-10-437-963-178103

Query Match      70.7%; Score 41; DB 4; Length 1114;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DWRFNAPFY 10
   |||||
Db 696 DWRFDVFOY 704

RESULT 5
US-10-361-522-3
; Sequence 3, Application US/10361522
; Publication No. US20030232406A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, John B
; APPLICANT: Sloma, Alan
; APPLICANT: Pero, Janice G
; APPLICANT: Hatch, Randolph T
; APPLICANT: Hermann, Theron
; APPLICANT: Erdenberger, Thomas
; TITLE OF INVENTION: BACTERIAL STRAINS WHICH OVERPRODUCE RIBOFLAVIN
; FILE REFERENCE: 13604 US6 (C38435/129129)
; CURRENT APPLICATION NUMBER: US/10/361.522
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: US 09/306,615
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1769
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-361-522-3

Query Match      69.0%; Score 40; DB 4; Length 1769;
Best Local Similarity 62.5%; Pred. No. 8.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDWRFNAPF 8
   |||||
Db 1660 FDMQYNAP 1667

RESULT 6
US-10-425-115-295441
; Sequence 295441, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 295441
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_32522C.1.pep
US-10-425-115-295441

Query Match      67.2%; Score 39; DB 4; Length 84;
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:17:02 ; Search time 8 Seconds  
(without alignments)  
16.403 Million cell updates/sec

Title: WRF

Perfect score: 58

Sequence: \*13ydwrfnafxy 10%

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA New:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	67.2	1368	6	US-10-770-303-2
2	39	67.2	1368	7	US-11-185-372-2
3	35	60.3	1476	6	US-10-647-956A-4
4	33	56.9	270	7	US-11-098-686-10898
5	32	55.2	362	7	US-11-012-762-56
6	32	55.2	415	7	US-11-203-109-20
7	32	55.2	550	7	US-11-098-686-10696
8	32	55.2	1152	7	US-11-080-025-4
9	31	53.4	39	6	US-10-895-064-295
10	31	53.4	126	7	US-11-064-174-10
11	31	53.4	126	7	US-11-064-174-144
12	31	53.4	126	7	US-11-064-174-150
13	31	53.4	127	7	US-11-064-174-11
14	31	53.4	163	7	US-11-186-284-57
15	31	53.4	201	6	US-10-467-657-5880
16	31	53.4	317	7	US-11-152-811-4
17	31	53.4	317	7	US-11-052-554A-48
18	31	53.4	552	6	US-10-131-826A-332
19	31	53.4	809	6	US-10-467-657-1106
20	31	53.4	810	7	US-11-052-554A-245
21	30	51.7	123	6	US-10-910-463-8
22	30	51.7	193	7	US-11-108-163B-14
23	30	51.7	201	7	US-11-170-653-23
24	30	51.7	217	7	US-11-108-163B-13
25	30	51.7	220	7	US-11-108-163B-12

26	30	51.7	289	7	US-11-052-554A-75	Sequence 75, Appl
27	30	51.7	301	7	US-11-108-163B-11	Sequence 11, Appl
28	30	51.7	344	7	US-11-108-163B-10	Sequence 10, Appl
29	30	51.7	429	6	US-10-886-504-3	Sequence 3, Appl
30	30	51.7	429	6	US-10-886-504-4	Sequence 4, Appl
31	30	51.7	429	6	US-10-886-505-3	Sequence 3, Appl
32	30	51.7	429	6	US-10-886-505-4	Sequence 4, Appl
33	30	51.7	429	6	US-10-886-527-3	Sequence 3, Appl
34	30	51.7	429	6	US-10-886-527-4	Sequence 4, Appl
35	30	51.7	436	7	US-11-080-991-102	Sequence 102, App
36	30	51.7	527	6	US-10-886-504-9	Sequence 9, Appl
37	30	51.7	527	6	US-10-886-504-10	Sequence 10, Appl
38	30	51.7	527	6	US-10-886-505-9	Sequence 9, Appl
39	30	51.7	527	6	US-10-886-505-10	Sequence 10, Appl
40	30	51.7	527	6	US-10-886-527-9	Sequence 9, Appl
41	30	51.7	527	6	US-10-886-527-10	Sequence 10, Appl
42	30	51.7	787	6	US-10-517-939-54	Sequence 54, Appl
43	30	51.7	1194	7	US-11-098-686-10270	Sequence 10270, A
44	30	51.7	2710	7	US-11-051-453-41	Sequence 41, Appl
45	29.5	50.9	2233	6	US-10-873-528-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-10-770-303-2  
; Sequence 2, Application US/10770303  
; Publication No. US20050283007A1  
; GENERAL INFORMATION:  
; APPLICANT: Uesugi, Montonari  
; APPLICANT: Asada, Shinichi  
; TITLE OF INVENTION: Small Molecule Inhibitors of Her2 Expression  
; FILE REFERENCE: HO-P02722US1/10213121  
; CURRENT APPLICATION NUMBER: US/10/770,303  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: US 60/380,481  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 10/405,387  
; PRIOR FILING DATE: 2004-01-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1368  
; TYPE: PRT  
; ORGANISM: Human  
US-10-770-303-2

Query Match 67.2%; Score 39; DB 6; Length 1368;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 DWRENAF 8  
Db 1091 DWRFNEF 1097

RESULT 2  
US-11-185-372-2  
; Sequence 2, Application US/11185372  
; Publication No. US20050288327A1  
; GENERAL INFORMATION:  
; APPLICANT: Uesugi, Montonari  
; APPLICANT: Asada, Shinichi  
; TITLE OF INVENTION: Small Molecule Inhibitors of Her2 Expression  
; FILE REFERENCE: HO-P02722US0/10213121  
; CURRENT APPLICATION NUMBER: US/11/185,372  
; CURRENT FILING DATE: 2005-07-20  
; PRIOR APPLICATION NUMBER: US 60/380,481  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

```
; LENGTH: 1368
; TYPE: PRT
; ORGANISM: Human
US-11-185-372-2

Query Match      67.2%; Score 39; DB 7; Length 1368;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DWRFNAF 8
Db 1091 DWRFNEF 1097

RESULT 3
US-10-647-956A-4
; Sequence 4, Application US/10647956A
; Publication No. US20050251878A1
; GENERAL INFORMATION:
; APPLICANT: Iffrench-Constant, Richard
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/10/647,956A
; CURRENT FILING DATE: 2003-08-26
; PRIOR FILING DATE: CURRENT FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1476
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-10-647-956A-4

Query Match      60.3%; Score 35; DB 6; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDWRFP 5
Db 1185 YDWRFP 1189

RESULT 4
US-11-098-686-10898
; Sequence 10898, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10898
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10898

Query Match      56.9%; Score 33; DB 7; Length 270;
Best Local Similarity 43.8%; Pred. No. 61;
```

```
Matches 7; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 YDWRFP-----NAF 8
Db 83 YEWRFYTIDDDTVNAF 98

RESULT 5
US-11-012-762-56
; Sequence 56, Application US/11012762
; Publication No. US20050244815A1
; GENERAL INFORMATION:
; APPLICANT: Georgia State University Research Foundation, Inc.
; TITLE OF INVENTION: Compositions and Methods for Viral Resistance Genes
; FILE REFERENCE: GSU1.PCT
; CURRENT APPLICATION NUMBER: US/11/012,762
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: PCT/US03/19300
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/390,046
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-012-762-56

Query Match      55.2%; Score 32; DB 7; Length 362;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDWRFNAP 8
Db 158 YAWRYNKY 165

RESULT 6
US-11-205-109-20
; Sequence 20, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-2US
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1)
; OTHER INFORMATION: V represents a non-standard initiator codon. It is expected that
; OTHER INFORMATION: the biosynthesized protein will have a formylmethionine residue
; OTHER INFORMATION: at this position
US-11-205-109-20

Query Match      55.2%; Score 32; DB 7; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WRFNA 7
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:08:07 ; Search time 19 Seconds  
(without alignments)  
50.640 Million cell updates/sec

Title: WRF  
Perfect score: 58  
Sequence: 1.ydwrfnafxy.1.0

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	72.4	561	2 G83913	hypothetical prote
2	38	65.5	363	2 G81445	hypothetical prote
3	37	63.8	233	2 B95218	hypothetical prote
4	37	63.8	233	2 A98082	UDGglucose 4-epime
5	37	63.8	643	2 C87412	ABC transporter, H
6	37	63.8	651	2 PC1123	hypothetical prote
7	37	63.8	811	2 E97040	phage related prot
8	37	63.8	1436	2 A99115	putative U5 snRNP-
9	36	62.1	125	2 E75624	hypothetical prote
10	36	62.1	327	2 D70346	conserved hypothet
11	36	62.1	499	2 C70345	phosphoglucumutase
12	36	62.1	566	2 JH0218	cellulase (EC 3.2.
13	36	62.1	680	2 H90093	guanine nucleotide
14	36	62.1	1483	2 E86143	F6F3.12 protein -
15	35	60.3	163	2 E81971	hypothetical prote
16	35	60.3	191	2 A97074	hypothetical prote
17	35	60.3	206	2 D86772	hypothetical prote
18	35	60.3	263	2 E97189	glycosyltransferas
19	35	60.3	268	2 B97189	glycosyltransferas
20	35	60.3	337	2 F70960	hypothetical prote
21	35	60.3	345	2 T25580	hypothetical prote
22	35	60.3	371	2 AC0115	conserved hypothet
23	35	60.3	408	2 F83379	conserved hypothet
24	35	60.3	418	2 H97763	proline/betaine tr
25	35	60.3	582	2 T05020	L-ascorbate oxidas
26	35	60.3	710	2 G82689	soluble lytic mure
27	35	60.3	720	2 T02734	hypothetical prote
28	35	60.3	929	2 T28927	hypothetical prote
29	34.5	59.5	797	2 T39643	probable cleavage

30	34	58.6	167	2 B71518	probable signal pe
31	34	58.6	322	2 E84908	hypothetical prote
32	34	58.6	332	2 S40762	hypothetical prote
33	34	58.6	333	2 T02690	hypothetical prote
34	34	58.6	335	2 B71361	probable thiamin A
35	34	58.6	357	2 E75284	conserved hypothet
36	34	58.6	380	2 I38435	angiotensin recept
37	34	58.6	383	2 T50502	hypothetical prote
38	34	58.6	407	2 H87341	OmpA family protei
39	34	58.6	456	2 F47677	phosphomannomutase
40	34	58.6	456	2 A50768	phosphomannomutase
41	34	58.6	456	2 H85830	phosphomannomutase
42	34	58.6	456	2 E90985	phosphomannomutase
43	34	58.6	456	2 B55239	phosphomannomutase
44	34	58.6	456	2 C90983	phosphomannomutase
45	34	58.6	456	2 F85828	phosphomannomutase

ALIGNMENTS

RESULT 1

G83913  
hypothetical protein BH2111 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C:Accession: G83913  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83913  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-561 <STO>  
A:Cross-references: UNIPROT:Q9XB24; UNIPARC:UPI000000C3DDA; GB:AP001514; GB:BA000004; NID  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2111

Query Match 72.4%; Score 42; DB 2; Length 561;  
Best Local Similarity 87.5%; Pred. No. 9.9; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy	1	YDWRFNAP	8
Db	227	YDWRFFAF	234

RESULT 2

G81445  
hypothetical protein Cj0272 [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: G81445  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrel  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: G81445  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-363 <PAR>  
A:Cross-references: UNIPROT:Q9PIM3; UNIPARC:UPI000000C20B2; GB:AL139074; GB:AL111168; NID  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0272  
C:Superfamily: Campylobacter jejuni hypothetical protein Cj0272

Query Match 65.5%; Score 38; DB 2; Length 363;  
Best Local Similarity 100.0%; Pred. No. 31; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0;

```

Qy      3 WRFNAP 8
      |||||
Db     291 WRFNAP 296

RESULT 3
B95218
hypothetical protein SPI867 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95218
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95218
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <KUR>
A:Cross-references: UNIPROT:Q97NY4; UNIPARC:UPI0000051A08; GB:AE005672; PIDN:AAK75939.1;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI867

Query Match      63.8%; Score 37; DB 2; Length 233;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 2; Indels 3; Gaps 0;

Qy      1 YDWRFNAPXY 10
      |||||
Db     162 YDKYVIFRY 171

RESULT 4
A98082
UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98082
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, C.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A98082
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <KUR>
A:Cross-references: UNIPROT:Q8DNJ6; UNIPARC:UPI000000E36D0; GB:AE007317; PIDN:AAI00486.1;
C:Genetics:
A:Gene: GalE-truncation
C:Keywords: isomerase

Query Match      63.8%; Score 37; DB 2; Length 233;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 2; Indels 3; Gaps 0;

Qy      1 YDWRFNAPXY 10
      |||||
Db     162 YDKYVIFRY 171

RESULT 5
C87412
ABC transporter, HlyB/MsbA family CC1314 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

```

```

C:Accession: C87412
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87412
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-643 <STO>
A:Cross-references: UNIPROT:Q9A8N9; UNIPARC:UPI00000C7343; GB:AE005673; NID:gl3422655; P
C:Genetics:
A:Gene: CC1314

Query Match      63.8%; Score 37; DB 2; Length 643;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YDWRFNAP 7
      |||||
Db     216 YDWRFAA 222

RESULT 6
PC1123
hypothetical protein - bloodfluke planorb (fragment)
C:Species: Biomphalaria glabrata (bloodfluke planorb)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Mar-1999
C:Accession: PC1123
R:Knight, M.; Miller, A.; Raghavan, N.; Richards, C.; Lewis, F.
Gene 118, 181-187, 1992
A:Title: Identification of a repetitive element in the snail Biomphalaria glabrata: Rela
A:Reference number: PC1123; MUID:92380502; PMID:1380940
A:Accession: PC1123
A:Molecule type: DNA
A:Residues: 1-651 <KNI>
A:Cross-references: UNIPARC:UPI000017BDA5; EMBL:X60372

Query Match      63.8%; Score 37; DB 2; Length 651;
Best Local Similarity 62.5%; Pred. No. 86;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YDWRFNAP 8
      |||||
Db     10 YDWRISF 17

RESULT 7
E97040
phage related protein, YonO B. subtilis homolog [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: E97040
R:Nolling, J.; Berton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Baly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E97040
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-811 <KUR>
A:Cross-references: UNIPROT:Q97JY3; UNIPARC:UPI000000CA0D9; GB:AE001437; PIDN:AAK79112.1;
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1139

Query Match      63.8%; Score 37; DB 2; Length 811;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 YDWRFNAP 8

```

## RESULT 2

```

Q6BCJ7_TETTH
ID Q6BCJ7_TETTH PRELIMINARY; PRT; 1275 AA.
AC Q6BCJ7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Dicer-related RNase III protein Dcllp.
GN Name=DCL1;
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymenidae; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=11599893; DOI=10.1101/gad.1265105;
RA Mochizuki K., Gorovsky M.A.;
RT "A Dicer-like protein in Tetrahymena has distinct functions in genome
rearrangement, chromosome segregation, and meiotic prophase.";
RL Genes Dev. 15:77-89(2005).
DR EMBL; AB182481; BAD34724.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR000999; RNase III.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS0142; RNASE_3_2; 2.
SQ SEQUENCE 1275 AA; 150259 MW; 6C7470D27980D55B CRC64;

Query Match 79.3%; Score 46; DB 2; Length 1275;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDWRFNAP 8
Db 816 YDWRFNLF 823

RESULT 3
Q813V5_PLAF7
ID Q813V5_PLAF7 PRELIMINARY; PRT; 2133 AA.
AC Q813V5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphatidylinositol 3-kinase, putative (EC 2.7.1.137).
GN Name=PF0765w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
Sulston J.E., Craig A., Newbold C., Barrall B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).

```

```

[2]
RN NUCLEOTIDE SEQUENCE.
RP Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
RA Hall N., Bowman S., Churcher C., Quail M., Barrall B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL929352; CAD51517.1; -; Genomic_DNA.
DR HSSP; P48736; 188Y.
DR GO; GO:0004428; F:inositol or phosphatidylinositol kinase act. . .; IEA.
DR GO; GO:0016303; F:phosphatidylinositol 3-kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000403; PI3/4_kinase_cat.
DR InterPro; IPR001263; PI3Ka.
DR Pfam; PF00454; PI3_P14_kinase; 1.
DR Pfam; PF00613; PI3Ka; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3KC; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS0290; PI3_4_KINASE_3; 1.
KW Kinase; Nucleotide-binding; Transferase.
SQ SEQUENCE 2133 AA; 255921 MW; CB2E6C4B196FA4FC CRC64;

Query Match 79.3%; Score 46; DB 2; Length 2133;
Best Local Similarity 70.0%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDWRFNAPXY 10
Db 774 YDWRFNAPXY 783

RESULT 4
Q9KB24_BACHD
ID Q9KB24_BACHD PRELIMINARY; PRT; 561 AA.
AC Q9KB24;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BH2111 protein.
GN OrderedLocNames=BH2111;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; BA000004; BAB05830.1; -; Genomic_DNA.
DR PIR; G83913; G83913.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
KW Complete proteome.
SQ SEQUENCE 561 AA; 64072 MW; BB76B1CA931D95A0 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 561;
Best Local Similarity 87.5%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDWRFNAP 8
Db 227 YDWRFNAP 234

RESULT 5
Q7PYW5_ANOGA
ID Q7PYW5_ANOGA PRELIMINARY; PRT; 479 AA.

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:01:37 ; Search time 116 Seconds  
(without alignments)  
37.877 Million cell updates/sec

Title: FRW

Perfect score: 58

Sequence: ~~(1-ydfeawafxy)10~~

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_21.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	98.3	10	9	Adw04632 Human AGR
2	57	98.3	10	9	Adw04631 Human AGR
3	46.5	80.2	11	9	Adw04668 Human AGR
4	46.5	80.2	11	9	Adw04657 Human AGR
5	46.5	80.2	11	9	Adw04654 Human AGR
6	46.5	80.2	11	9	Adw04653 Human AGR
7	46.5	80.2	11	9	Adw04649 Human AGR
8	46.5	80.2	11	9	Adw04655 Human AGR
9	46.5	80.2	11	9	Adw04634 Human AGR
10	46.5	80.2	11	9	Adw04635 Human AGR
11	43.5	75.0	11	9	Adw04662 Human AGR
12	42	72.4	10	9	Adw04630 Human AGR
13	42	72.4	10	9	Adw04629 Human AGR
14	40.5	69.8	11	9	Adw04651 Human AGR
15	40	69.0	499	2	Aaw55660 H. pylori
16	40	69.0	528	2	Aaw55732 H. pylori
17	40	69.0	528	2	Aaw58295 H. pylori
18	40	69.0	528	2	Aay17167 H. pylori
19	40	69.0	528	5	Abb07348 H. pylori
20	40	69.0	528	5	Abb07349 H. pylori
21	40	69.0	528	5	Abb07347 H. pylori
22	40	69.0	528	6	Abu30791 Protein e
23	40	69.0	528	9	Aeb91407 Microbial
24	40	69.0	881	6	Abu23568 Protein e

25	39.5	68.1	11	9	Adw04661	Adw04661 Human AGR
26	39.5	68.1	11	9	Adw04660	Adw04660 Human AGR
27	39.5	68.1	11	9	Adw04650	Adw04650 Human AGR
28	39.5	68.1	11	9	Adw04656	Adw04656 Human AGR
29	39.5	68.1	11	9	Adw04658	Adw04658 Human AGR
30	39.5	68.1	11	9	Adw04659	Adw04659 Human AGR
31	39	67.2	874	6	Abu48867	Protein e
32	38	65.5	310	4	AAU29393	Human G p
33	38	65.5	310	5	ABG60681	Novel G p
34	38	65.5	484	8	ADx80139	Plant ful
35	37	63.8	260	6	ABM69356	Phototrab
36	37	63.8	302	3	AAy97213	Campyloba
37	37	63.8	302	6	ABJ18492	Campyloba
38	37	63.8	303	6	ABU26594	Protein e
39	37	63.8	309	3	AAAB18898	A maize c
40	37	63.8	416	8	ADU05585	H. pylori
41	36	62.1	18	8	ADP47038	Murine he
42	36	62.1	128	8	ADP46948	Murine he
43	36	62.1	274	5	ABP64898	Human pro
44	36	62.1	274	7	ADD68864	Human ABC
45	36	62.1	323	2	AAAR35084	(Pro)leuk

#### ALIGNMENTS

##### RESULT 1

Adw04632

ID Adw04632 standard; peptide; 10 AA.

XX AC Adw04632;

XX DT 10-MAR-2005 (first entry)

XX DE Human AGR/MCR agonist chimeric cyclic peptide #6.

XX KW protein engineering; melanocortin receptor; AGRP; agouti related protein; obesity; gene therapy; anorectic; cyclic; melanocortin agonist.

XX OS Homo sapiens.

XX OS Chimeric.

XX OS Unidentified.

XX FH Key

FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at position 9"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 9 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a lactam bridge with Asp at position 2"

FT Modified-site 10 /note= "C-terminal amide"

XX US2004260063-A1.

XX PD 23-DEC-2004.

XX PF 23-JUN-2003; 2003US-00602394.

XX PR 23-JUN-2003; 2003US-00602394.

XX (HASK/) HASKELL-LUEVANO C.

XX PI Haskell-Luevano C;

XX XX WPI; 2005-030733/03.

XX PT New chimeric peptides and templates based upon melanocortin agonist peptides and agouti related protein antagonist peptide, useful for treating or preventing conditions modulated by melanocortin receptors, such as obesity.

XX PS Claim 2; SEQ ID NO 7; 15pp; English.

XX CC The present invention relates to a chimeric peptide that is biologically

XX CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related

XX CC protein (AGRP) template and melanocortin agonist-based bioactive

XX CC determinant sequences which have been substituted for the analogous

XX CC template sequences. The invention is useful for treating or preventing

XX CC various diseases and conditions modulated by melanocortin receptors, such

XX CC as obesity and in gene therapy. The present sequence is the human

XX CC AGRP/MCR agonist chimeric cyclic peptide.

XX SQ Sequence 10 AA;

Query Match 98.3%; Score 57; DB 9; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.008;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDFRWNAFY 10

Db 1 YDFRWNAFY 10

|||||

RESULT 2

ADW04631

ID ADW04631 standard; peptide; 10 AA.

XX AC ADW04631;

DT 10-MAR-2005 (first entry)

XX DE Human AGRP/MCR agonist chimeric cyclic peptide #5.

XX KW protein engineering; melanocortin receptor; AGRP; agouti related protein;

XX KW obesity; gene therapy; anorectic; cyclic; melanocortin agonist.

XX OS Homo sapiens.

XX OS Chimeric.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at

FT position 9"

FT Misc-difference 9

FT /note= "2,3-Diaminopropionic acid (Dpr) residue forms a

FT lactam bridge with Asp at position 2"

XX PN US2004260063-A1.

XX PD 23-DEC-2004.

XX PF 23-JUN-2003; 2003US-00602394.

XX PR 23-JUN-2003; 2003US-00602394.

XX PA (HASK/) HASKELL-LUEVANO C.

XX PI Haskell-Luevano C;

XX DR WPI; 2005-030733/03.

XX PS New chimeric peptides and templates based upon melanocortin agonist

XX PS peptides and agouti related protein antagonist peptide, useful for

XX PT treating or preventing conditions modulated by melanocortin receptors,

XX PT such as obesity.

XX PS Claim 2; SEQ ID NO 6; 15pp; English.

XX CC The present invention relates to a chimeric peptide that is biologically

XX CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related

XX CC protein (AGRP) template and melanocortin agonist-based bioactive

XX CC determinant sequences which have been substituted for the analogous

XX CC template sequences. The invention is useful for treating or preventing

XX CC various diseases and conditions modulated by melanocortin receptors, such

XX CC as obesity and in gene therapy. The present sequence is the human

XX CC AGRP/MCR agonist chimeric cyclic peptide.

XX SQ Sequence 10 AA;

Query Match 98.3%; Score 57; DB 9; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.008;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDFRWNAFY 10

Db 1 YDFRWNAFY 10

|||||

RESULT 3

ADW04668

ID ADW04668 standard; peptide; 11 AA.

XX AC ADW04668;

DT 10-MAR-2005 (first entry)

XX DE Human AGRP/MCR agonist chimeric cyclic peptide #29.

XX KW protein engineering; melanocortin receptor; AGRP; agouti related protein;

XX KW obesity; gene therapy; anorectic; melanocortin agonist; cyclic.

XX OS Homo sapiens.

XX OS Chimeric.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at

FT position 10"

FT Misc-difference 4 /label= D-form residue

FT Misc-difference 10

FT /note= "2,3-Diaminopropionic acid (Dpr) residue forms a

FT lactam bridge with Asp at position 2"

XX PN US2004260063-A1.

XX PD 23-DEC-2004.

XX PF 23-JUN-2003; 2003US-00602394.

XX PR 23-JUN-2003; 2003US-00602394.

XX PA (HASK/) HASKELL-LUEVANO C.

XX PI Haskell-Luevano C;

XX DR WPI; 2005-030733/03.

XX PS New chimeric peptides and templates based upon melanocortin agonist

XX PS peptides and agouti related protein antagonist peptide, useful for

XX PT treating or preventing conditions modulated by melanocortin receptors,

XX PT such as obesity.

XX PS Claim 5; SEQ ID NO 43; 15pp; English.

XX CC The present invention relates to a chimeric peptide that is biologically

XX CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related

XX CC protein (AGRP) template and melanocortin agonist-based bioactive

XX CC determinant sequences which have been substituted for the analogous

XX CC template sequences. The invention is useful for treating or preventing

XX CC various diseases and conditions modulated by melanocortin receptors, such

XX CC as obesity and in gene therapy. The present sequence is the human

XX CC AGRP/MCR agonist chimeric cyclic peptide.



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

# OM protein - protein search, using sw model

Run on: February 14, 2006, 20:14:30 ; Search time 30.6667 Seconds  
(without alignments)  
26.959 Million cell updates/sec

Title: FRW  
Perfect score: 58  
Sequence: 1 ydfrwnafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

- Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCITUS-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	63.8	303	2	US-09-495-406-17
2	37	63.8	303	2	US-09-816-028A-29
3	37	63.8	303	2	US-10-303-162-29
4	37	63.8	303	2	US-10-303-134-29
5	37	63.8	303	2	US-10-303-118-29
6	37	63.8	303	2	US-10-303-128-29
7	37	63.8	309	2	US-09-522-714-10
8	36	62.1	675	2	US-09-489-039A-9046
9	36	62.1	967	2	US-09-489-039A-8766
10	36	62.1	975	2	US-09-543-681A-5755
11	35	60.3	16	1	US-08-637-759B-344
12	35	60.3	16	2	US-08-871-355A-344
13	35	60.3	16	2	US-09-201-945-344
14	35	60.3	305	2	US-09-635-872A-3
15	35	60.3	305	2	US-09-636-077A-3
16	35	60.3	305	2	US-09-636-060C-3
17	35	60.3	305	2	US-09-636-552-3
18	35	60.3	305	2	US-09-636-596C-3
19	35	60.3	305	2	US-10-023-894-7
20	35	60.3	305	2	US-10-306-686-3
21	35	60.3	305	2	US-09-895-072-3
22	35	60.3	305	2	US-10-023-888-7
23	35	60.3	307	2	US-09-635-872A-9
24	35	60.3	307	2	US-09-636-077A-9
25	35	60.3	307	2	US-09-636-060C-9
26	35	60.3	307	2	US-09-586-552-9
27	35	60.3	307	2	US-09-636-596C-9

28	35	60.3	307	2	US-10-023-894-12	Sequence 12, Appl
29	35	60.3	307	2	US-10-306-686-9	Sequence 9, Appl
30	35	60.3	307	2	US-09-895-072-9	Sequence 9, Appl
31	35	60.3	307	2	US-10-023-888-12	Sequence 12, Appl
32	35	60.3	452	2	US-09-563-794B-140	Sequence 140, Appl
33	35	60.3	961	2	US-09-328-352-4182	Sequence 4182, Ap
34	34	58.6	84	2	US-09-270-767-58928	Sequence 58928, A
35	34	58.6	86	2	US-09-270-767-40527	Sequence 40527, A
36	34	58.6	86	2	US-09-270-767-55743	Sequence 55743, A
37	34	58.6	190	2	US-09-543-681A-7684	Sequence 7684, Ap
38	34	58.6	206	2	US-09-248-796A-17898	Sequence 17898, A
39	34	58.6	208	2	US-09-489-039A-10423	Sequence 10423, A
40	34	58.6	208	2	US-09-393-171-1	Sequence 1, Appl
41	34	58.6	226	2	US-09-270-767-41679	Sequence 41679, A
42	34	58.6	301	2	US-09-134-001C-4305	Sequence 4305, Ap
43	34	58.6	332	2	US-09-710-279-2350	Sequence 2350, Ap
44	34	58.6	390	2	US-09-634-238-331	Sequence 331, App
45	34	58.6	390	2	US-09-489-039A-8316	Sequence 8316, Ap

## ALIGNMENTS

RESULT 1  
US-09-495-406-17  
; Sequence 17, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
US-09-495-406-17

Query Match 63.8%; Score 37; DB 2; Length 303;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 YDFRWNAFY 10  
| | | | |  
Db 290 YQFKWKXFLY 299

RESULT 2  
US-09-816-028A-29  
; Sequence 29, Application US/09816028A  
; Patent No. 669705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49

—

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:15:26 ; Search time 127 Seconds  
(without alignments)  
32.900 Million cell updates/sec

Title: FRW  
Perfect score: 58  
Sequence: 1 ydfrwnafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2\_6/ptodata/1/pubpaa/us07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/us08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/us09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/us10\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/us10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/us11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	69.0	77	US-10-335-977-4929	Sequence 4929, Ap
2	40	69.0	499	US-10-335-977-4932	Sequence 4932, Ap
3	40	69.0	528	US-09-881-752A-190	Sequence 190, App
4	40	69.0	528	US-10-282-122A-58715	Sequence 58715, A
5	40	69.0	528	US-10-335-977-4933	Sequence 4933, Ap
6	40	69.0	528	US-10-662-126-31	Sequence 31, Appl
7	40	69.0	881	US-10-282-122A-51492	Sequence 51492, A
8	39	67.2	874	US-10-282-122A-76791	Sequence 76791, A
9	38	65.5	310	US-09-811-284-142	Sequence 142, App
10	38	65.5	484	US-10-425-114-49505	Sequence 49505, A
11	38	65.5	728	US-10-424-599-189465	Sequence 189465, A
12	37	63.8	159	US-10-335-977-8565	Sequence 8565, Ap
13	37	63.8	205	US-10-424-599-145298	Sequence 145298, A
14	37	63.8	273	US-10-156-761-8617	Sequence 8617, Ap
15	37	63.8	303	US-09-816-028A-29	Sequence 29, Appl
16	37	63.8	303	US-10-303-161-29	Sequence 29, Appl
17	37	63.8	303	US-10-303-118-29	Sequence 29, Appl
18	37	63.8	303	US-10-303-128-29	Sequence 29, Appl
19	37	63.8	303	US-10-303-134-29	Sequence 29, Appl
20	37	63.8	303	US-10-303-162-29	Sequence 29, Appl
21	37	63.8	303	US-10-282-122A-54518	Sequence 54518, A
22	37	63.8	303	US-10-820-536-29	Sequence 29, Appl
23	37	63.8	303	US-10-845-408-29	Sequence 29, Appl
24	37	63.8	303	US-10-845-412-29	Sequence 29, Appl
25	37	63.8	303	US-10-846-219-29	Sequence 29, Appl
26	37	63.8	303	US-10-821-604-29	Sequence 29, Appl
27	37	63.8	303	US-10-847-983-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1

US-10-335-977-4929  
; Sequence 4929, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/335,977  
; FILING DATE: 30-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,002  
; FILING DATE: 17-DEC-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: GTN-018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 4929:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 77 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...77  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4929:  
US-10-335-977-4929

28	37	63.8	303	5	US-10-821-573-29	Sequence 29, Appl
29	37	63.8	303	5	US-10-850-807-29	Sequence 29, Appl
30	37	63.8	303	5	US-10-850-125-29	Sequence 29, Appl
31	37	63.8	303	5	US-10-830-825-29	Sequence 29, Appl
32	37	63.8	303	5	US-10-962-334-29	Sequence 29, Appl
33	37	63.8	303	5	US-10-830-997-29	Sequence 29, Appl
34	37	63.8	303	5	US-10-962-235-29	Sequence 29, Appl
35	37	63.8	303	5	US-10-961-882-29	Sequence 29, Appl
36	37	63.8	307	4	US-10-767-701-47078	Sequence 47078, A
37	37	63.8	309	4	US-10-304-928-10	Sequence 323462, A
38	37	63.8	313	4	US-10-425-115-323462	Sequence 11624, A
39	37	63.8	391	5	US-10-732-923-11624	Sequence 133410, A
40	37	63.8	567	4	US-10-437-963-133410	Sequence 133413, A
41	37	63.8	773	4	US-10-437-963-133410	Sequence 94, Appl
42	36	62.1	18	5	US-10-725-962-94	Sequence 4, Appl
43	36	62.1	128	5	US-10-725-962-4	Sequence 174643, A
44	36	62.1	135	4	US-10-424-599-174643	Sequence 59, Appl
45	36	62.1	254	6	US-11-090-847-59	

Query Match 69.0%; Score 40; DB 4; Length 77;  
 Best Local Similarity 60.0%; Pred. No. 47;  
 Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 YDFRWNAFY 10  
 DB 12 YDWRNNAEY 21

## RESULT 2

US-10-335-977-4932  
 ; Sequence 4932, Application US/10335977  
 ; Publication No. US20040052799A1  
 ; GENERAL INFORMATION:  
 ; TITLE OF INVENTION: DOUGLAS SMITH et al  
 ; RELATING TO HELICOBACTER PYLORI FOR  
 ; DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 10031  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 28 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109-1875  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: IBM PC Compatible  
 ; OPERATING SYSTEM: Windows NT 4.0  
 ; SOFTWARE: UNIX  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/335,977  
 ; FILING DATE: 30-Dec-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/993,002  
 ; FILING DATE: 17-DEC-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mandragouras, Amy E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: GTN-018  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617)227-7400  
 ; TELEFAX: (617)742-4214  
 ; INFORMATION FOR SEQ ID NO: 4932:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 499 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Helicobacter pylori  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (B) LOCATION 1...499  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4932:  
 US-10-335-977-4932

Query Match 69.0%; Score 40; DB 4; Length 499;  
 Best Local Similarity 60.0%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YDFRWNAFY 10  
 DB 308 YDWRNNAEY 317

## RESULT 3

US-09-881-752A-190  
 ; Sequence 190, Application US/09881752A  
 ; Patent No. US20020115078A1  
 ; GENERAL INFORMATION:

; APPLICANT: Kleanthous, Harold  
 ; APPLICANT: Al-Garawi, Amal  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Tomb, Jean-Francois  
 ; APPLICANT: Oomen, Raymond P.  
 ; TITLE OF INVENTION: Identification of Polynucleotides  
 ; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in the  
 ; TITLE OF INVENTION: Genome  
 ; FILE REFERENCE: 06132/041002  
 ; CURRENT APPLICATION NUMBER: US/09/881,752A  
 ; CURRENT FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 08/833,457  
 ; PRIOR FILING DATE: 1997-04-01  
 ; NUMBER OF SEQ ID NOS: 370  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 190  
 ; LENGTH: 528  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori  
 US-09-881-752A-190

Query Match 69.0%; Score 40; DB 3; Length 528;  
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YDFRWNAFY 10  
 DB 308 YDWRNNAEY 317

## RESULT 4

US-10-282-122A-58715  
 ; Sequence 58715, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:17:02 ; Search time 8 Seconds  
(without alignments)  
16.403 Million cell updates/sec

Title: FRW  
Perfect score: 58  
Sequence: 1 ydfrwnafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US03\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	69.0	528	7	US-11-052-554A-117
2	36	62.1	454	7	US-11-024-959-458
3	36	62.1	550	7	US-11-098-686-10696
4	34	58.6	332	6	US-10-793-626-2350
5	34	58.6	483	6	US-10-793-626-3132
6	33	56.9	294	7	US-11-098-686-10545
7	33	56.9	518	7	US-11-024-959-364
8	33	56.9	529	7	US-11-037-243-62
9	33	56.9	556	6	US-10-063-703-24
10	33	56.9	556	6	US-11-102-240-24
11	33	56.9	603	6	US-10-793-626-1684
12	33	56.9	625	7	US-11-055-822-522
13	33	56.9	756	6	US-10-131-826A-392
14	32.5	56.0	363	6	US-10-517-939-264
15	32	55.2	93	6	US-10-467-657-5984
16	32	55.2	143	6	US-10-793-626-370
17	32	55.2	269	6	US-10-467-657-2198
18	32	55.2	335	6	US-10-873-528-4
19	32	55.2	396	6	US-11-032-773-941
20	32	55.2	548	6	US-10-793-626-336
21	32	55.2	599	7	US-11-165-141-33
22	32	55.2	614	7	US-11-015-546A-20
23	32	55.2	622	7	US-11-098-686-11191
24	32	55.2	651	6	US-10-994-820A-34
25	32	55.2	763	7	US-11-013-247A-35

Sequence 23, Appl  
Sequence 95, Appl  
Sequence 110, Appl  
Sequence 11343, A  
Sequence 10686, A  
Sequence 33, Appl  
Sequence 36, Appl  
Sequence 41, Appl  
Sequence 6, Appl  
Sequence 306, Appl  
Sequence 332, Appl  
Sequence 520, Appl  
Sequence 4, Appl  
Sequence 304, Appl  
Sequence 330, Appl  
Sequence 518, Appl  
Sequence 44, Appl  
Sequence 43, Appl  
Sequence 10893, A

26 32 55.2 1075 7 US-11-089-551A-23  
27 32 55.2 1237 7 US-11-052-554A-95  
28 32 55.2 1420 7 US-11-077-550-110  
29 31.5 54.3 693 7 US-11-098-686-11343  
30 31 53.4 234 7 US-11-098-686-10686  
31 31 53.4 338 7 US-11-207-626A-33  
32 31 53.4 341 7 US-11-207-626A-36  
33 31 53.4 343 7 US-11-207-626A-40  
34 31 53.4 393 7 US-11-207-626A-41  
35 31 53.4 410 7 US-11-055-822-6  
36 31 53.4 410 7 US-11-055-822-306  
37 31 53.4 410 7 US-11-055-822-332  
38 31 53.4 410 7 US-11-055-822-520  
39 31 53.4 456 7 US-11-055-822-4  
40 31 53.4 456 7 US-11-055-822-304  
41 31 53.4 456 7 US-11-055-822-330  
42 31 53.4 456 7 US-11-055-822-518  
43 31 53.4 520 7 US-11-094-917-44  
44 31 53.4 526 7 US-11-094-917-43  
45 31 53.4 882 7 US-11-098-686-10893

ALIGNMENTS

RESULT 1  
US-11-052-554A-117  
; Sequence 117, Application US/11052554A  
; Publication No. US2005028866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 117  
; LENGTH: 528  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori J99  
US-11-052-554A-117

Query Match 69.0%; Score 40; DB 7; Length 528;  
Best Local Similarity 60.0%; Pred. No. 5.7;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YDFRWNAFX 10  
Db 308 YDFRWNAFX 317

RESULT 2  
US-11-024-959-458  
; Sequence 458, Application US/11024959  
; Publication No. US20060010516A1  
; GENERAL INFORMATION:  
; APPLICANT: FORSTER, RICHARD L.  
; APPLICANT: CONNETT, SARAH JANE  
; APPLICANT: EMERSON, SARAH JANE  
; APPLICANT: GRIGOR, MURRAY ROBERT  
; APPLICANT: HIGGINS, COLLEEN M.  
; APPLICANT: LUND, STEVEN TROY  
; APPLICANT: MAGUSIN, ANDREAS  
; APPLICANT: KODRZYCKI, BOB  
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS  
; FILE REFERENCE: 044463-0360  
; CURRENT APPLICATION NUMBER: US/11/024,959

```

; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 458
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-458

Query Match      62.1%; Score 36; DB 7; Length 454;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 DFRWNAF 8
Db      394 DFHNSP 400
|||||

RESULT 3
US-11-098-686-10696
; Sequence 10696, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR FILING DATE: 2002-10-04
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10696
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10696

Query Match      62.1%; Score 36; DB 7; Length 550;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 YDFRWNAFY 10
Db      427 YDFRNNAYEY 436
|||||

RESULT 4
US-10-793-626-2350
; Sequence 2350, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence

US-10-793-626-2350
Query Match      58.6%; Score 34; DB 6; Length 332;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YDFRWNA 7
Db      131 HDFRWEA 137
|||||

US-10-793-626-3132
; Sequence 3132, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3132
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence

US-10-793-626-3132
Query Match      58.6%; Score 34; DB 6; Length 483;
Best Local Similarity 62.5%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3 FRWNAFY 10
Db      332 FDWNSPKY 339
|||||

RESULT 6
US-11-098-686-10545
; Sequence 10545, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10545
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10545

Query Match      56.9%; Score 33; DB 7; Length 294;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 YDFRWNAF 8
Db      175 YDLQWNEY 182
|||||
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:08:07 ; Search time 19 Seconds  
(without alignments)  
50.640 Million cell updates/sec

Title: FRW  
Perfect score: 58  
Sequence: 1 ydfrwnafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	69.0	528	F64580	hypothetical prote
2	40	69.0	528	F71931	outer membrane pro
3	40	69.0	881	G87195	valyl-tRNA synthet
4	39	67.2	205	C84334	hypothetical prote
5	39	67.2	874	E82913	valyl-tRNA synthet
6	38	65.5	395	D69312	molybdopterin oxid
7	38	65.5	955	F84972	valine-tRNA ligase
8	37	63.8	276	B71336	hypothetical prote
9	37	63.8	303	F81318	probable galactosy
10	37	63.8	416	C64590	hypothetical prote
11	37	63.8	416	D71923	hypothetical prote
12	37	63.8	426	D71367	conserved hypoteth
13	37	63.8	660	E87391	TonB-dependent rec
14	36	62.1	300	B49238	gamma-hemolysin ga
15	36	62.1	323	JN0627	leukocidin chain F
16	36	62.1	323	JQ1530	leukocidin chain F
17	36	62.1	325	F90043	gamma-hemolysin co
18	36	62.1	325	B49234	leucocidin R, comp
19	36	62.1	325	JC7909	ATP-binding casset
20	36	62.1	531	T45865	hypothetical prote
21	36	62.1	664	T33379	hypothetical prote
22	36	62.1	951	1 SEECVT	valine-tRNA ligase
23	36	62.1	951	AC1061	valine-tRNA ligase
24	36	62.1	951	C91283	valine-tRNA synthet
25	36	62.1	951	E86124	valine-tRNA synthet
26	36	62.1	954	E64121	valine-tRNA ligase
27	36	62.1	956	B71250	valine-tRNA ligase
28	36	62.1	965	B50418	valine-tRNA ligase
29	36	62.1	994	B82843	valyl-tRNA synthet

30	35	60.3	64	2	A97790	hypothetical prote
31	35	60.3	267	2	AE3589	3-oxoadipate enol-
32	35	60.3	269	2	T00088	ABC-type transport
33	35	60.3	305	2	T45062	hypothetical prote
34	35	60.3	320	2	T21839	hypothetical prote
35	35	60.3	379	2	T31154	hypothetical prote
36	35	60.3	408	2	B90517	conserved hypoteth
37	35	60.3	452	2	S72266	translation initia
38	35	60.3	465	2	T41511	probable dolichol
39	35	60.3	533	2	H69326	conserved hypoteth
40	35	60.3	745	2	S44792	F09G8.8 protein -
41	35	60.3	956	2	T23570	hypothetical prote
42	35	60.3	1211	2	D64116	exodeoxyribonuclea
43	35	60.3	1483	2	E86143	F6F3.12 protein -
44	34.5	59.5	355	2	AF1608	branched-chain fat
45	34	58.6	118	1	PSBGA	phospholipase A2 (

ALIGNMENTS

RESULT 1

F64580  
hypothetical protein HP0486 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: F64580  
R:Tomb, J.F.; White, O.; Kervlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.W.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.; A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: F64580  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-528 <TOM>  
A:Cross-references: UNIPROT:O25230; UNIPARC:UPI00000C084E; GB:AE000563; GB:AE000511; NID C:Superfamily: Helicobacter pylori hypothetical protein HP0209

Query Match 69.0%; Score 40; DB 2; Length 528;  
Best Local Similarity 60.0%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YDFRWNAFXY 10  
||:|:|  
Db 308 YDFRWNAEY 317

RESULT 2

F71931  
outer membrane protein - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: F71931  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: F71931  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-528 <ARN>  
A:Cross-references: UNIPROT:Q9ZLY7; UNIPARC:UPI00000D3626; GB:AE001478; GB:AE001439; NID C:Experimental source: strain J99

A:Gene: jhp0438  
C:Superfamily: Helicobacter pylori hypothetical protein HP0209

Query Match 69.0%; Score 40; DB 2; Length 528;

```

Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YDFRWNAFY 10
Db 308 YDFRWNAEY 317
||| |||
||| |||

RESULT 3
G97195
valyl-tRNA synthetase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97195
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-881 <KUR>
A;Cross-references: UNIPROT:Q97GG8; UNIPARC:UPI000000CA4EA; GB:AE001437; PIDN:AAK80354.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
C;Superfamily: valine-tRNA ligase
A;Gene: CAC2399

Query Match 69.0%; Score 40; DB 2; Length 881;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDFRWNAF 8
Db 641 YDFWMNEF 648
||| |||
||| |||

RESULT 4
C84334
hypothetical protein Vng183lh [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84334
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
.; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabic
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <STO>
A;Cross-references: UNIPROT:Q9HP31; UNIPARC:UPI0000006399E; GB:AE004437; NID:gl0581277; F
C;Genetics:
A;Gene: VNG1831H

Query Match 67.2%; Score 39; DB 2; Length 205;
Best Local Similarity 85.7%; Pred. No. 8.6;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDFRWNA 7
Db 192 YDIRWNA 198
||| |||
||| |||

RESULT 5
B82913
valyl-tRNA synthetase UU267 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

```

```

C;Accession: B82913
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: B82913
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-874 <GLA>
A;Cross-references: UNIPARC:UPI000013670D; GB:AE002123; GB:AF222894; NID:G6899229; PIDN:
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: vals; UU267
A;Genetic code: SGC3
C;Superfamily: valine-tRNA ligase

Query Match 67.2%; Score 39; DB 2; Length 874;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDFRWNAF 8
Db 637 YDFIWNTF 644
||| |||
||| |||

RESULT 6
D69312
molybdopterin oxidoreductase, membrane subunit homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69312
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Goeyne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: D69312
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-395 <KLE>
A;Cross-references: UNIPROT:O29750; UNIPARC:UPI00000570A9; GB:AE001069; GB:AE000782; NID

Query Match 65.5%; Score 38; DB 2; Length 395;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FRWNAFY 10
Db 127 FSWNAFLY 134
||| |||
||| |||

RESULT 7
F84972
valine-tRNA ligase (EC 6.1.1.9) [imported] - Buchnera sp. (strain APS)
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C;Accession: F84972
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: F84972
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-955 <STO>
A;Cross-references: UNIPARC:UPI000005E561; GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: vals; BU366
C;Superfamily: valine-tRNA ligase

```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:01:55 ; Search time 126.333 Seconds  
(without alignment)  
55.847 Million cell updates/sec

Title: FRW  
Perfect score: 58  
Sequence: 1 ydfrnafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	75.9	858	2	Q86IY5 dictyosteli
2	41	70.7	513	2	Q8XKG7 CLOPE
3	41	70.7	944	2	Q4UQP3_XANCP
4	41	70.7	944	2	Q8PCR7_XANCP
5	41	70.7	944	2	Q8PGQ7_XANAC
6	41	70.7	980	2	Q5H4N5_XANOR
7	41	70.7	1263	2	Q50SH1_ENTHI
8	40	69.0	293	2	Q7PMK7_ANOGA
9	40	69.0	469	2	Q744R2_MYCPA
10	40	69.0	528	2	Q25230_HELPY
11	40	69.0	528	2	Q8ZLH7_HELPY
12	40	69.0	880	2	Q891R5_CLOTE
13	40	69.0	881	2	Q97G98_CLOAB
14	40	69.0	937	2	Q4U920_THEAN
15	40	69.0	954	2	Q4N1R5_THEPA
16	39	67.2	205	2	Q8HP31_HALSA
17	39	67.2	480	2	Q4I2V9_GIBBEZ
18	39	67.2	560	2	Q4I0D0_GIBBEZ
19	39	67.2	590	2	Q5P0V3_AZOSE
20	39	67.2	874	1	SVU_UREPA
21	39	67.2	960	1	SVU_BUCAP
22	39	67.2	1572	2	Q4XV99_PLACH
23	39	67.2	1616	2	Q4YU25_PLABE
24	39	67.2	1623	2	Q7PDV5_PLAYO
25	39	67.2	1844	2	Q37287_PLAF7
26	38	65.5	309	2	Q6AFG1_LEIXX
27	38	65.5	383	2	Q6CH91_YARLI
28	38	65.5	395	1	HNEB_ARCFU
29	38	65.5	556	2	Q5NW45_AZOSE
30	38	65.5	627	2	Q5BD44_EMENI
31	38	65.5	703	2	Q4WC74_ASPFU

Q8eup6 mycoplasma  
P57447 buchnera ap  
Q4ugr2 theileria a  
O7m558 brachydanio  
Q7n246 photorhabdu  
Q985r2 thizobium l  
Q82p55 streptomyce  
Q83367 treponema p  
P96089 thermoanaer  
Q9jrq2 thermoanaer  
Q9pnf5 campylobact  
Q4mj80 erratitia pr  
Q4tmj2 erythrobaet  
Q9hjz2 thermoplasma

## ALIGNMENTS

## RESULT 1

Q86IY5 D1CDI  
ID Q86IY5 D1CDI PRELIMINARY; PRT; 858 AA.  
AC Q86IY5;2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=DD80167886;  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;  
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A., Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G., Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;  
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";  
RL Nature 418:79-85(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RA Baumgart C.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A., Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q., Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F., Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P., Pilcher K., Chen G., Saunders D., Sodergren E., Davis P., Karhounou A., Nie X., Hall N., Anfar C., Hemphill L., Bason N., Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C., Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I., Muzny D., Mourlet T., Pain A., Lu M., Harper D., Lindsay R., Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C., Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A., Louisedge H., Mungall K., Oliver K., Price C., Quail M.A., Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A., Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y., Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C., Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R., Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;  
RT "The genome of the social amoeba Dictyostelium discoideum.";  
RL Nature 0:0-0(2005).  
DR EMBL; AC116986; AAC051863.1; -; Genomic DNA.  
DR EMBL; AAF101000027; EAL70063.1; -; Genomic DNA.  
DR DictyBase; DDB0168226; JCV22\_0\_00892.  
DR InterPro; IPR011020; HTTM.

```

DR SMART; SM00752; HTTM; 1.
KW Hypothetical protein.
SQ SEQUENCE 858 AA; 98593 MW; 040A389DB900EC85 CRC64;

Query Match 75.9%; Score 44; DB 2; Length 858;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YDFRWNAFY 10
   ||: ||| ||
Db 706 YDYSWQFAY 715

RESULT 2
ID Q8XKG7 CLOPE PRELIMINARY; PRT; 513 AA.
AC Q8XKG7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CPE1429.
GN OrderedLocusNames=CPE1429;
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; BA000016; BAB81135.1; -; Genomic DNA.
DR GO; GO:0015969; P:guanosine tetraphosphate metabolism; IEA.
DR InterPro; IPR007685; RelA_Spot.
DR Pfam; PF04607; RelA_Spot; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 513 AA; 60477 MW; 159D0E05B4A0B1EA CRC64;

Query Match 70.7%; Score 41; DB 2; Length 513;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YDFRWNAF 8
   ||| ||| ||
Db 335 YDFHWNLF 342

RESULT 3
ID Q4UQP3 XANCP PRELIMINARY; PRT; 944 AA.
AC Q4UQP3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Valyl-tRNA synthetase.
GN ORFNames=XC3588;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

```

```

CC -!- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC -!- + L-valyl-tRNA(Val).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR EMBL; CP000050; AAY50630.1; -; Genomic_DNA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_I.
DR InterPro; IPR002303; tRNA-synt_val.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00986; TRNASYNTHAL.
DR TIGRFAMs; TIGR00422; vals; 1.
DR PROSITE; PS00178; AA tRNA LIGASE I; 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
KW Protein biosynthesis.
SQ SEQUENCE 944 AA; 106095 MW; 73EA22380436B10D CRC64;

Query Match 70.7%; Score 41; DB 2; Length 944;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YDFRWNAF 8
   ||: ||| |||
Db 699 YEFANNAF 706

RESULT 4
ID Q8PCR7 XANCP PRELIMINARY; PRT; 944 AA.
AC Q8PCR7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Valyl-tRNA synthetase.
GN Names=vals;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Ciapina L.E.A., Camarotte G., Cannavaro F., Cardoso J., Chambergo F.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Scubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012162; AAM39962.1; -; Genomic_DNA.
DR HSP; P96142; IIVS
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004832; P:valine-tRNA ligase activity; IEA.
DR GO; GO:0006438; P:valyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_I.
DR InterPro; IPR002303; tRNA-synt_val.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00986; TRNASYNTHAL.
DR TIGRFAMs; TIGR00422; vals; 1.
DR PROSITE; PS00178; AA tRNA LIGASE I; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome.

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: February 14, 2006, 20:01:37 ; Search time 116 Seconds  
(without alignments)  
37.877 Million cell updates/sec  
Title: HFRW  
Perfect score: 60  
Sequence: ~~h~~-ydhfrwafxy-10  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : A\_Geneseq\_21:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*  
9: Geneseqp2005s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	80.8	11	9	ADW04668 Human AGR
2	48.5	80.8	11	9	ADW04657 Human AGR
3	48.5	80.8	11	9	ADW04634 Human AGR
4	48.5	80.8	11	9	ADW04635 Human AGR
5	45.5	75.8	11	9	ADW04662 Human AGR
6	42.5	70.8	11	9	ADW04651 Human AGR
7	41.5	69.2	11	9	ADW04661 Human AGR
8	41.5	69.2	11	9	ADW04660 Human AGR
9	41.5	69.2	11	9	ADW04658 Human AGR
10	41.5	69.2	11	9	ADW04659 Human AGR
11	40.5	67.5	11	9	ADW04650 Human AGR
12	40	66.7	8	4	ABB80033 Template
13	40	66.7	89	2	AAy23670 Human bet
14	40	66.7	424	8	ADN24572 Bacterial
15	40	66.7	431	8	Adq79875 T. thermo
16	40	66.7	431	8	Adq79873 T. thermo
17	40	66.7	431	8	Adq79877 T. thermo
18	40	66.7	431	8	Adq79881 Thermus s
19	40	66.7	431	8	Adq79700 Thermus t
20	40	66.7	431	8	Adq79706 Thermus t
21	40	66.7	431	8	Adq79698 Thermus t
22	40	66.7	431	8	Adq79702 Thermus f
23	40	66.7	431	8	Adq75571 Thb beta-
24	40	66.7	431	8	Adq75569 Tat beta-

25	40	66.7	431	8	ADQ75573 Tfi beta-
26	40	66.7	431	8	Adq75577 Tib10 bet
27	40	66.7	436	8	Adq79879 Thermus s
28	40	66.7	436	8	Adq79704 Thermus t
29	40	66.7	436	8	Adq75575 Tib2 beta
30	40	66.7	449	8	AdS30418 Bacterial
31	40	66.7	451	8	Adn21812 Bacterial
32	40	66.7	488	8	Adn25473 Bacterial
33	40	66.7	513	2	AAW42119 Coniferin
34	40	66.7	556	6	ABU21343 Protein e
35	39.5	65.8	11	9	Adw04654 Human AGR
36	39.5	65.8	11	9	Adw04655 Human AGR
37	39	65.0	6	3	AAb11848 Alpha-MSH
38	39	65.0	7	3	AAb11852 Alpha-MSH
39	39	65.0	8	3	AAb11856 Alpha-MSH
40	38.5	64.2	11	9	Adw04656 Human AGR
41	38.5	64.2	11	9	Adw04653 Human AGR
42	38.5	64.2	11	9	Adw04649 Human AGR
43	38	63.3	161	6	ABM69386 PhotOrhab
44	38	63.3	395	2	AAy37052 Protein i
45	38	63.3	660	6	ABU27245 Protein e

ALIGNMENTS

RESULT 1  
ADW04668  
ID ADW04668 standard; peptide; 11 AA.  
XX  
AC ADW04668;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Human AGRP\MCR agonist chimeric cyclic peptide #29.  
XX  
KW protein engineering; melanocortin receptor; AGRP; agouti related protein;  
KW obesity; gene therapy; anorectic; melanocortin agonist; cyclic.  
XX  
OS Homo sapiens.  
OS Chimeric.  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at position 10"  
FT Misc-difference 4 /label= D-form residue  
FT Misc-difference 10 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a lactam bridge with Asp at position 2"  
PN US2004260063-A1.  
XX 23-DEC-2004.  
XX 23-JUN-2003; 2003US-00602394.  
XX 23-JUN-2003; 2003US-00602394.  
XX (HASK/) HASKELL-LUEVANO C.  
XX Haskell-Luevano C;  
XX WPI; 2005-030733/03.  
XX New chimeric peptides and templates based upon melanocortin agonist peptides and agouti related protein antagonist peptide, useful for treating or preventing conditions modulated by melanocortin receptors, such as obesity.  
XX Claim 5; SEQ ID NO 43; 15pp; English.  
PS

CC The present invention relates to a chimeric peptide that is biologically  
 CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.  
 SQ Sequence 11 AA;

Query Match 80.8%; Score 48.5; DB 9; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 YDHFWR-AEYX 10  
 DB 1 YDHFWRNAFYX 11

RESULT 2  
 ADW04657  
 ID ADW04657 standard; peptide; 11 AA.  
 AC ADW04657;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX Human AGRP\MCR agonist chimeric cyclic peptide #18.  
 DE protein engineering; melanocortin receptor; AGRP; agouti related protein;  
 XX obesity; gene therapy; anorectic; melanocortin agonist; cyclic.  
 KW Homo sapiens.  
 OS Chimeric.  
 OS Unidentified.  
 XX

Key Location/Qualifiers  
 FH Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at  
 FT position 10"  
 FT Modified-site 4 /note= "D-form residue, (pI)Dpne"  
 FT Misc-difference 10 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a  
 FT lactam bridge with Asp at position 2"  
 XX US2004260063-A1.  
 XX 23-DEC-2004.  
 XX 23-JUN-2003; 2003US-00602394.  
 XX 23-JUN-2003; 2003US-00602394.  
 XX (HASK/) HASKELL-LUEVANO C.  
 XX Haskell-Luevano C;  
 XX WPI; 2005-030733/03.  
 XX New chimeric peptides and templates based upon melanocortin agonist  
 XX peptides and agouti related protein antagonist peptide, useful for  
 XX treating or preventing conditions modulated by melanocortin receptors,  
 XX such as obesity.  
 XX Claim 5; SEQ ID NO 32; 15pp; English.  
 XX The present invention relates to a chimeric peptide that is biologically  
 CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous

CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.  
 SQ Sequence 11 AA;

Query Match 80.8%; Score 48.5; DB 9; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 YDHFWR-AEYX 10  
 DB 1 YDHFWRNAFYX 11

RESULT 3  
 ADW04634  
 ID ADW04634 standard; peptide; 11 AA.  
 AC ADW04634;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX Human AGRP\MCR agonist chimeric cyclic peptide #8.  
 DE protein engineering; melanocortin receptor; AGRP; agouti related protein;  
 KW obesity; gene therapy; anorectic; cyclic; melanocortin agonist.  
 XX Homo sapiens.  
 OS Chimeric.  
 OS Unidentified.  
 XX

Key Location/Qualifiers  
 FH Misc-difference 2 /note= "This residue forms a lactam bridge with Dpr at  
 FT position 10"  
 FT Misc-difference 10 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a  
 FT lactam bridge with Asp at position 2"  
 XX US2004260063-A1.  
 XX 23-DEC-2004.  
 XX 23-JUN-2003; 2003US-00602394.  
 XX 23-JUN-2003; 2003US-00602394.  
 XX (HASK/) HASKELL-LUEVANO C.  
 XX Haskell-Luevano C;  
 XX WPI; 2005-030733/03.  
 XX New chimeric peptides and templates based upon melanocortin agonist  
 XX peptides and agouti related protein antagonist peptide, useful for  
 XX treating or preventing conditions modulated by melanocortin receptors,  
 XX such as obesity.  
 XX Claim 2; SEQ ID NO 9; 15pp; English.  
 XX The present invention relates to a chimeric peptide that is biologically  
 CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.  
 SQ Sequence 11 AA;

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

# OM protein - protein search, using sw model

Run on: February 14, 2006, 20:14:30 ; Search time 30.6667 Seconds  
(without alignments)  
26.959 Million cell updates/sec

Title: HFRW

Perfect score: 60

Sequence: <1 ydhrwfafxy 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5-COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/6-COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/H-COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	66.7	89	2	US-09-217-228-29
2	40	66.7	513	1	US-09-122-230-7
3	38	63.3	662	2	US-09-198-452A-169
4	38	63.3	665	2	US-09-438-185A-151
5	37	61.7	14	1	US-08-480-190-176
6	37	61.7	14	1	US-08-488-379-176
7	37	61.7	14	2	US-08-475-399A-176
8	37	61.7	14	2	US-08-077-255A-176
9	37	61.7	14	4	PCT-US93-07545-176
10	37	61.7	396	2	US-08-860-656B-3
11	37	61.7	396	2	US-08-860-656B-4
12	37	61.7	441	2	US-08-914-375C-63
13	37	61.7	448	2	US-09-306-593-12
14	37	61.7	497	2	US-09-902-540-15207
15	37	61.7	1836	2	US-10-162-012-24
16	37	61.7	6	2	US-09-623-548A-34
17	36	60.0	6	2	US-09-657-276-34
18	36	60.0	6	2	US-08-349-902B-15
19	36	60.0	7	1	US-08-240-078-54
20	36	60.0	7	2	US-09-374-827-6
21	36	60.0	7	2	US-09-953-349-6
22	36	60.0	7	2	US-09-458-580-6
23	36	60.0	105	2	US-08-851-362D-30
24	36	60.0	136	2	US-09-621-976-3913
25	36	60.0	162	1	US-08-624-125-8
26	36	60.0	162	2	US-08-937-155-8
27	36	60.0	162	2	US-08-937-155-8

28	36	60.0	162	2	US-09-323-998E-8
29	36	60.0	355	1	US-09-014-969-19
30	36	60.0	477	2	US-09-991-181-285
31	36	60.0	477	2	US-09-990-444-285
32	36	60.0	477	2	US-09-997-333-285
33	36	60.0	477	2	US-09-992-598-285
34	36	60.0	567	1	US-08-504-459-2
35	36	60.0	567	1	US-08-504-459-4
36	36	60.0	567	1	US-08-504-459-6
37	35	58.3	86	2	US-09-270-767-36243
38	35	58.3	86	2	US-09-270-767-51460
39	35	58.3	439	2	US-08-914-375C-67
40	35	58.3	609	2	US-09-257-525A-9
41	35	58.3	671	2	US-09-843-598-5
42	35	58.3	671	2	US-09-843-598-7
43	35	58.3	1024	2	US-09-562-737-88
44	35	58.3	3418	1	US-08-639-501-2
45	35	58.3	3418	1	US-08-603-753D-4

## ALIGNMENTS

### RESULT 1

US-09-217-228-29

; Sequence 29, Application US/09217228

; Patent No. 6323178

; GENERAL INFORMATION:

; APPLICANT: Butler, Jon P.

; APPLICANT: Hale, John E.

; APPLICANT: Heath Jr., William F.

; APPLICANT: Schoner, Brigitte E.

; APPLICANT: Heiman, Mark L.

; APPLICANT: Becker, Gerald W.

; APPLICANT: Varshavsky, Alexander D.

; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof

; FILE REFERENCE: X-12139

; CURRENT APPLICATION NUMBER: US/09/217,228

; CURRENT FILING DATE: 1998-12-21

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 29

; LENGTH: 89

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Human analog

US-09-217-228-29

Query Match 66.7%; Score 40; DB 2; Length 89;

Best Local Similarity 100.0%; Pred. No. 7.4;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DHRWA 7

Db 46 DHRWA 51

### RESULT 2

US-09-122-230-7

; Sequence 7, Application US/09122230A

; Patent No. 5973228

; GENERAL INFORMATION:

; APPLICANT: Carlson, et al.

; TITLE OF INVENTION: Coniferin Beta Glucosidase cDNA for Modifying Lignin

; TITLE OF INVENTION: Composition in Plants

; FILE REFERENCE: 50532

; CURRENT APPLICATION NUMBER: US/09/122,230A

; CURRENT FILING DATE: 1998-07-23

; EARLIER APPLICATION NUMBER: U.S. 60/053,566

; EARLIER FILING DATE: 1997-07-24

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 7
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Pinus contorta
US-09-122-230-7

Query Match      66.7%; Score 40; DB 1; Length 513;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      2 DHRWAFXY 10
Db      461 DNFWAFGY 469

RESULT 3
US-09-198-452A-169
; Sequence 169, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 169
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-169

Query Match      63.3%; Score 38; DB 2; Length 662;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3 HFRWAFXY 10
Db      302 HYRWALAY 309

RESULT 4
US-09-438-185A-151
; Sequence 151, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 151
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0149
US-09-438-185A-151

Query Match      63.3%; Score 38; DB 2; Length 665;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;

; SEQ ID NO 1
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Roman M. Chicz
US-08-480-190-176

Query Match      61.7%; Score 37; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 3.8;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 YDHRWAF 8
Db      6 YDTFSWAF 13

RESULT 6
US-08-488-379-176
; Sequence 176, Application US/08488379
; Patent No. 5880103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:15:26 ; Search time 127 Seconds  
(without alignments)  
32.900 Million cell updates/sec

Title: HFRW

Perfect score: 60

Sequence: 1-ydhrfwafxy 10 \*

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	75.0	320	US-10-732-923-23387	Sequence 23387, A
2	42	70.0	688	US-10-425-115-198285	Sequence 198285, A
3	40	66.7	89	US-09-824-438-29	Sequence 29, Appl
4	40	66.7	207	US-10-437-963-144257	Sequence 144257, A
5	40	66.7	424	US-10-369-493-7225	Sequence 7225, Ap
6	40	66.7	449	US-10-369-493-19451	Sequence 19451, A
7	40	66.7	451	US-10-369-493-4465	Sequence 4465, Ap
8	40	66.7	488	US-10-369-493-8126	Sequence 8126, Ap
9	40	66.7	519	US-10-425-115-306500	Sequence 306500, A
10	40	66.7	556	US-10-282-122A-49267	Sequence 49267, A
11	38	63.3	345	US-10-732-923-3030	Sequence 3030, Ap
12	38	63.3	457	US-10-437-963-104080	Sequence 104080, Ap
13	38	63.3	660	US-10-282-122A-55169	Sequence 55169, A
14	38	63.3	662	US-10-289-762-169	Sequence 169, App
15	38	63.3	662	US-10-282-122A-54853	Sequence 54853, A
16	37	61.7	28	US-10-362-024A-24	Sequence 24, Appl
17	37	61.7	136	US-10-425-115-283784	Sequence 283784, A
18	37	61.7	229	US-11-097-143-33174	Sequence 33174, A
19	37	61.7	296	US-10-108-260A-3020	Sequence 3020, Ap
20	37	61.7	382	US-10-739-930-7183	Sequence 7183, Ap
21	37	61.7	413	US-10-276-774-2206	Sequence 2206, Ap
22	37	61.7	452	US-10-156-761-9340	Sequence 9340, Ap
23	37	61.7	457	US-10-156-761-10667	Sequence 10667, A
24	37	61.7	482	US-10-156-761-12787	Sequence 12787, A
25	37	61.7	779	US-10-282-122A-45594	Sequence 45594, A
26	37	61.7	786	US-10-282-122A-45859	Sequence 45859, A
27	37	61.7	1836	US-09-875-363-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-10-732-923-23387  
; Sequence 23387, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgarton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 23387  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
US-10-732-923-23387

Query Match 75.0%; Score 45; DB 5; Length 320;  
Best Local Similarity 75.0%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YDHRWAF 8  
DB 257 YDHRWFPY 264

RESULT 2

US-10-425-115-198285  
; Sequence 198285, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 198285  
; LENGTH: 688  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(688)

```
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112414C.1.pep
US-10-425-115-198285

Query Match          70.0%; Score 42; DB 4; Length 688;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDHFRA 7
    |||||
Db 314 YDFEWA 320

RESULT 3
US-09-824-438-29
; Sequence 29, Application US/09824438
; Publication No. US20030073621A1
; GENERAL INFORMATION:
; APPLICANT: Butler, John P.
; APPLICANT: Hale, John E.
; APPLICANT: Heath Jr., William F.
; APPLICANT: Schoner, Brigitte E.
; APPLICANT: Heiman, Mark L.
; APPLICANT: Becker, Gerald W.
; APPLICANT: Varshavsky, Alexander D.
; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof
; FILE REFERENCE: X-12139
; CURRENT APPLICATION NUMBER: US/09/824,438
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 29
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human analog
US-09-824-438-29

Query Match          66.7%; Score 40; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DHFRWA 7
    |||||
Db 46 DHFRWA 51

RESULT 4
US-10-437-963-144257
; Sequence 144257, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144257
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)..(207)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45089C.1.pep
US-10-437-963-144257

Query Match          66.7%; Score 40; DB 4; Length 207;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YDHFRAFX 10
    |||||
Db 89 YDVRMGFV 98

RESULT 5
US-10-369-493-7225
; Sequence 7225, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7225
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
; OTHER INFORMATION: Description of Artificial Sequence: Human analog
US-10-369-493-7225

Query Match          66.7%; Score 40; DB 4; Length 424;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DHFRWAFX 10
    |||||
Db 395 DNFWAFX 403

RESULT 6
US-10-369-493-19451
; Sequence 19451, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19451
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; OTHER INFORMATION: Description of Artificial Sequence: Human analog
US-10-369-493-19451
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:17:02 ; Search time 8 Seconds  
(without alignments)  
16.403 Million cell updates/sec

Title: HFRW  
Perfect score: 60  
Sequence: Q1 ydhrwfafxy 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 97014 segs, 13122538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pap.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pap.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pap.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	60.0	349	7	US-11-098-686-10281
2	36	60.0	477	6	US-10-131-826A-452
3	35	58.3	396	6	US-10-467-657-280
4	35	58.3	396	6	US-10-467-657-4380
5	35	58.3	1111	7	US-11-142-700-26
6	34	56.7	15	6	US-10-939-890-198
7	34	56.7	336	7	US-11-224-260-3
8	33	55.0	188	7	US-11-234-786-592
9	33	55.0	375	6	US-10-517-939-2
10	33	55.0	422	7	US-11-230-251-26
11	33	55.0	777	7	US-11-043-693-3
12	33	55.0	1298	7	US-11-076-427A-12
13	33	55.0	1362	7	US-11-043-693-33
14	33	55.0	1363	7	US-11-043-693-32
15	33	55.0	1363	7	US-11-076-427A-10
16	33	55.0	1368	7	US-11-043-693-34
17	33	55.0	1981	6	US-10-374-954-23
18	33	55.0	1998	6	US-10-374-954-21
19	33	55.0	2009	6	US-10-374-954-2
20	32	53.3	5	7	US-11-019-955-6
21	32	53.3	13	7	US-11-019-955-4
22	32	53.3	13	7	US-11-188-552-1
23	32	53.3	13	7	US-11-188-552-71
24	32	53.3	14	7	US-11-188-552-2
25	32	53.3	14	7	US-11-188-552-3

26	32	53.3	14	7	US-11-188-552-4	Sequence 4, Appli
27	32	53.3	14	7	US-11-188-552-5	Sequence 5, Appli
28	32	53.3	14	7	US-11-188-552-6	Sequence 6, Appli
29	32	53.3	14	7	US-11-188-552-11	Sequence 11, Appl
30	32	53.3	14	7	US-11-188-552-12	Sequence 12, Appl
31	32	53.3	14	7	US-11-188-552-13	Sequence 13, Appl
32	32	53.3	17	7	US-11-083-418-4	Sequence 4, Appli
33	32	53.3	70	7	US-11-019-955-8	Sequence 8, Appli
34	32	53.3	169	7	US-11-129-143-183	Sequence 183, App
35	32	53.3	241	7	US-11-019-955-27	Sequence 27, Appl
36	32	53.3	241	7	US-11-019-955-29	Sequence 29, Appl
37	32	53.3	263	7	US-11-112-882-35	Sequence 35, Appl
38	32	53.3	263	7	US-11-166-993-120	Sequence 120, App
39	32	53.3	268	7	US-11-019-955-28	Sequence 28, Appl
40	32	53.3	315	6	US-10-524-647-25	Sequence 25, Appl
41	32	53.3	320	6	US-10-524-647-4	Sequence 4, Appli
42	32	53.3	329	6	US-10-524-647-2	Sequence 2, Appli
43	32	53.3	329	6	US-10-524-647-23	Sequence 23, Appl
44	32	53.3	329	6	US-10-524-647-80	Sequence 80, Appl
45	32	53.3	341	6	US-10-524-647-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-11-098-686-10281  
; Sequence 10281, Application US/11098686  
; Publication No. US2006024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: 09531-128001  
; CURRENT APPLICATION NUMBER: US/11/098.686  
; PRIOR FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: PCT/US03/31318  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: US 60/416,395  
; NUMBER OF SEQ ID NOS: 11433  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10281  
; LENGTH: 349  
; TYPE: PRT  
; ORGANISM: Lawsonia intracellularis  
US-11-098-686-10281

Query Match 60.0%; Score 36; DB 7; Length 349;  
Best Local Similarity 62.5%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YDHRWFAP 8  
Db 40 YTHFNWKP 47

RESULT 2  
US-10-131-826A-452  
; Sequence 452, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven

```

; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131.826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 452
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-452

```

```

Query Match          60.0%; Score 36; DB 6; Length 477;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 YDHRWAFXY 10
Db 35 PDGFRWDYLY 44

```

```

RESULT 3
; Sequence 280, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 280
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-280

Query Match          58.3%; Score 35; DB 6; Length 396;

```

```

Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDHRWAF 8
Db 111 YDRFRFAF 118

```

```

RESULT 4
US-10-467-657-4380
; Sequence 4380, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4380
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4380

```

```

Query Match          58.3%; Score 35; DB 6; Length 396;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDHRWAF 8
Db 111 YDRFRFAF 118

```

```

RESULT 5
US-11-142-700-26
; Sequence 26, Application US/11142700
; Publication No. US20060026721A1
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Gary M. Fader
; APPLICANT: Saverio Carl Falco
; APPLICANT: Anthony J. Kinney
; APPLICANT: Jonathan E. Lightner
; APPLICANT: Guo-Hua Miao
; APPLICANT: J. Antoni Rafaleski
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB-1170
; CURRENT APPLICATION NUMBER: US/11/142,700
; CURRENT FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: US/09/720,383
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 1111
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-142-700-26

```

```

Query Match          58.3%; Score 35; DB 7; Length 1111;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2006, 20:08:07 ; Search time 19 Seconds  
(without alignments)  
50.640 Million cell updates/sec

Title: HFRW

Perfect score: 60

Sequence: -1 ydhfrwafxy 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_80:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	66.7	460	J5137	beta-glucosidase (
2	39	65.0	186	T29117	hypothetical prote
3	39	65.0	320	T16358	hypothetical prote
4	39	65.0	537	T41666	carl homolog - fis
5	38	63.3	345	T07777	probable homeobox
6	38	63.3	660	A81704	DNA ligase TC0423
7	38	63.3	662	B72114	DNA ligase CP0624
8	38	63.3	662	E86509	DNA ligase (import
9	38	63.3	663	C71551	probable DNA ligas
10	37	61.7	164	G84716	hypothetical prote
11	37	61.7	311	T29111	probable beta-giuc
12	37	61.7	448	S17215	beta-glucosidase (
13	37	61.7	473	A48949	beta-glucosidase,
14	37	61.7	660	A84232	spore cortex synth
15	37	61.7	786	H69980	single-strand DNA-
16	37	61.7	1835	I34323	sodium channel alp
17	37	61.7	1836	I64893	sodium channel alp
18	37	61.7	1836	J50648	sodium channel alp
19	37	61.7	1836	I51964	sodium channel alp
20	37	61.7	1840	CHRTM1	sodium channel pro
21	37	61.7	1920	S43721	lactase (EC 3.2.1.
22	36	60.0	57	T27749	hypothetical prote
23	36	60.0	117	H71108	hypothetical prote
24	36	60.0	131	A84724	hypothetical prote
25	36	60.0	165	S27679	hypothetical prote
26	36	60.0	225	S28679	glycoprotein 30 -
27	36	60.0	324	E85513	probable integrase
28	36	60.0	324	G90662	probable integrase
29	36	60.0	408	T38386	hypothetical wd-40

30	36	60.0	434	2	S51644	secreted/adhesive
31	36	60.0	466	2	S44878	ZC262.6 protein -
32	36	60.0	477	1	A59390	probable phosphodi
33	36	60.0	477	1	A59391	probable phosphodi
34	36	60.0	477	2	JC8005	nucleotide pyropho
35	36	60.0	525	2	T32481	hypothetical prote
36	36	60.0	567	2	S69779	adhesin AP65-2 pre
37	36	60.0	567	2	S69778	adhesin AP65-1 pre
38	36	60.0	1181	2	D86157	hypothetical prote
39	36	60.0	1204	2	S62506	alpha-glucan synth
40	36	60.0	1820	1	CHEE	sodium channel pro
41	36	60.0	2049	2	T43161	sodium channel pro
42	35.5	59.2	190	2	F82810	conserved hypothe
43	35	58.3	74	2	E44103	neurocalcin gamma
44	35	58.3	148	2	F71207	hypothetical prote
45	35	58.3	166	2	AD3624	hypothetical prote

ALIGNMENTS

RESULT 1

JC5137

beta-glucosidase (EC 3.2.1.21) - Bifidobacterium breve

N:Alternate names: beta-D-glucosidase

N:Contains: beta-D-fucosidase (EC 3.2.1.38); beta-galactosidase (EC 3.2.1.23)

C:Species: Bifidobacterium breve

C:Date: 31-Jan-1997 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004

C:Accession: JC5137; PC4240

R:Nunoura, N.; Ohdan, K.; Tanaka, K.; Tanaki, H.; Yano, T.; Inui, M.; Yukawa, H.; Yamano

Biosci. Biotechnol. Biochem. 60, 2011-2018, 1996

A:Title: Cloning and nucleotide sequence of the beta-D-glucosidase gene from Bifidobacte

A:Reference number: JC5137; MUID:97142514; PMID:8988633

A:Accession: JC5137

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-460 <NNUN1>

A:Cross-references: UNIPROT:P94248; UNIPARC:UPI00000826ED; DDBJ:D88311; DDBJ:D84489; NID

A:Accession: PC4240

A>Status: preliminary

A:Molecule type: protein

A:Residues: 2-29 <NNUN2>

A:Cross-references: UNIPARC:UPI0000175B1B

C:Comment: This enzyme also has beta-D-fucosidase and beta-D-galactosidase activities.

C:Superfamily: Agrobacterium beta-glucosidase

C:Keywords: glycosidase; hydrolase

F:2-460/Product: beta-glucosidase #status predicted <MAT>

F:119,168/Active site: His, Glu #status predicted

Query Match 66.7%; Score 40; DB 2; Length 460;

Best Local Similarity 66.7%; Pred. No. 15;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DHFRWAFXY 10

Db 417 DNFEWAFGY 425

RESULT 2

T29117

hypothetical protein SCIF2.08c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29117

R:Parkhill, J.; Barrall, B. G.; Rajandream, M. A.

submitted to the EMBL Data Library, August 1998

A:Reference number: Z17215

A:Accession: T29117

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-186 <PAR>

A:Cross-references: UNIPROT:O86545; UNIPARC:UPI00000DAD8D; EMBL:AL031350; NID:el316892;

C:Genetics:

A;Note: SCIF2.08c

Query Match 65.0%; Score 39; DB 2; Length 186;  
Best Local Similarity 75.0%; Pred. No. 9;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDHRWAF 8  
| | | | |  
Db 179 YSHTRWAF 186

RESULT 3  
Tl6358  
hypothetical protein f43c9.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: Tl6358  
R;Fulton, B.  
submitted to the EMBL Data Library, November 1995  
A;Description: The sequence of C. elegans cosmid f43c9.  
A;Reference number: Z18499  
A;Accession: Tl6358  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-320 <FUL>  
A;Cross-references: UNIPROT:Q20359; UNIPARC:UPI00000761D4; EMBL:U40427; NID:gl065557; PI  
C;Genetics:  
A;Gene: CESP:f43c9.1  
A;Introns: 32/3; 75/3; 126/2; 155/3; 231/3; 289/3  
C;Superfamily: Caenorhabditis elegans hypothetical protein f43c9.1

Query Match 65.0%; Score 39; DB 2; Length 320;  
Best Local Similarity 83.3%; Pred. No. 16;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDHRW 6  
| | | | |  
Db 178 FDHRW 183

RESULT 4  
T41666  
carl homolog - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T41666  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.  
submitted to the EMBL Data Library, May 1998  
A;Reference number: Z22008  
A;Accession: T41666  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-537 <WOO>  
A;Cross-references: UNIPROT:O59833; UNIPARC:UPI000006BBEA; EMBL:AL023590; PIDN:CAA15073.  
A;Experimental source: strain 972h-; cosmid c965  
C;Genetics:  
A;Gene: SPDB:SPCC965.13  
A;Map position: 3  
C;Superfamily: benomyl/methotrexate resistance protein

Query Match 65.0%; Score 39; DB 2; Length 537;  
Best Local Similarity 62.5%; Pred. No. 27;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDHRWAF 8  
| | | | |  
Db 238 YDHRWFY 245

RESULT 5  
T07777  
probable homeobox protein H1 - potato  
N;Alternate names: POTH1

C;Species: Solanum tuberosum (potato)  
C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 05-Oct-2004  
C;Accession: T07777  
R;Hart, J.K.; Hannapel, D.J.  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z16126  
A;Accession: T07777  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-345 <HAR>  
A;Cross-references: UNIPROT:P91588; UNIPARC:UPI00000A31E3; EMBL:U65648; NID:gl814233; PI  
A;Experimental source: cv. Superior; grown under short days; developmental stage: 4-day  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;262-325/Region: homeobox

Query Match 63.3%; Score 38; DB 2; Length 345;  
Best Local Similarity 71.4%; Pred. No. 26;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DHRWAF 8  
| | | | |  
Db 284 DHRWPFY 290

RESULT 6  
A81704  
DNA ligase TC0423 [imported] - Chlamydia muridarum (strain Nigg)  
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 31-Dec-2004  
C;Accession: A81704  
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A;Reference number: A81500; MUID:20150255; PMID:10684935  
A;Accession: A81704  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-660 <TET>  
A;Cross-references: UNIPROT:Q9PKP2; UNIPARC:UPI0000057903; GB:AE002309; GB:AE002160; NID  
A;Experimental source: strain Nigg (MoPn)  
C;Genetics:  
A;Gene: TC0423  
C;Superfamily: DNA ligase (NAD), Liga type

Query Match 63.3%; Score 38; DB 2; Length 660;  
Best Local Similarity 62.5%; Pred. No. 51;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 HFRWAFXY 10  
| : | | | |  
Db 301 HYRWAIAY 308

RESULT 7  
B72114  
DNA ligase CP0624 [imported] - Chlamydophila pneumoniae (strains CWL029 and AR39)  
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 31-Dec-2004  
C;Accession: B72114; C81556  
R;Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A;Reference number: A72000; MUID:99206606; PMID:10192388  
A;Accession: B72114  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-662 <ARN>  
A;Cross-references: UNIPROT:Q9Z934; UNIPARC:UPI000012967D; GB:AE001601; GB:AE001363; NID  
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
  
OM protein - protein search, using sw model  
  
Run on: February 14, 2006, 20:01:55 ; Search time 126.333 Seconds  
(without alignments)  
55.847 Million cell updates/sec  
  
Title: HFRW  
Perfect score: 60  
Sequence: @1-ydhfrwaixy 10f

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
  
Searched: 2166443 seqs, 705528306 residues  
  
Total number of hits satisfying chosen parameters: 2166443  
  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES											
Result	No.	Score	Query Match	Length	DB ID	Description					
1	45	75.0	320	2	Q9ZRC0	LYCES	Q9ZRC0 lycopersico	Q9ZRC0	LYCES	Q9ZRC0	LYCES
2	40	66.7	30	2	Q7RNH7	PLAYO	Q7RNH7 plasmodium	Q7RNH7	PLAYO	Q7RNH7	PLAYO
3	40	66.7	255	2	Q53KQ3	ORYSA	Q53KQ3 oryza sativ	Q53KQ3	ORYSA	Q53KQ3	ORYSA
4	40	66.7	392	2	Q7NKW9	GLOVI	Q7NKW9 gloeobacter	Q7NKW9	GLOVI	Q7NKW9	GLOVI
5	40	66.7	423	2	Q9A588	9DEIN	Q9A588 thermus sp.	Q9A588	9DEIN	Q9A588	9DEIN
6	40	66.7	427	2	Q94721	ORYSA	Q94721 oryza sativ	Q94721	ORYSA	Q94721	ORYSA
7	40	66.7	431	2	Q8GEB0	9DEIN	Q8GEB0 thermus sp.	Q8GEB0	9DEIN	Q8GEB0	9DEIN
8	40	66.7	431	2	Q8GEB2	THEFI	Q8GEB2 thermus fil	Q8GEB2	THEFI	Q8GEB2	THEFI
9	40	66.7	431	2	Q8GEB3	THETH	Q8GEB3 thermus the	Q8GEB3	THETH	Q8GEB3	THETH
10	40	66.7	431	2	Q8GEB4	THETH	Q8GEB4 thermus the	Q8GEB4	THETH	Q8GEB4	THETH
11	40	66.7	431	2	Q8GEB5	THECA	Q8GEB5 thermus cal	Q8GEB5	THECA	Q8GEB5	THECA
12	40	66.7	431	2	Q9A611	THETH	Q9A611 thermus the	Q9A611	THETH	Q9A611	THETH
13	40	66.7	431	2	Q53W75	THET8	Q53W75 thermus the	Q53W75	THET8	Q53W75	THET8
14	40	66.7	431	2	Q74611	THET2	Q74611 thermus the	Q74611	THET2	Q74611	THET2
15	40	66.7	436	2	Q9L794	9DEIN	Q9L794 thermus non	Q9L794	9DEIN	Q9L794	9DEIN
16	40	66.7	436	2	Q8GEB1	9DEIN	Q8GEB1 thermus sp.	Q8GEB1	9DEIN	Q8GEB1	9DEIN
17	40	66.7	460	2	P94248	BIFBR	P94248 bifidobacte	P94248	BIFBR	P94248	BIFBR
18	40	66.7	513	2	Q9ZT64	PINCO	Q9ZT64 pinus conto	Q9ZT64	PINCO	Q9ZT64	PINCO
19	40	66.7	557	2	Q61029	CAEBR	Q61029 caenorhabdi	Q61029	CAEBR	Q61029	CAEBR
20	40	66.7	748	2	Q61M36	CAEBR	Q61M36 caenorhabdi	Q61M36	CAEBR	Q61M36	CAEBR
21	39	65.0	186	2	Q86545	STRCO	Q86545 streptomyce	Q86545	STRCO	Q86545	STRCO
22	39	65.0	320	2	Q20359	CAEEL	Q20359 caenorhabdi	Q20359	CAEEL	Q20359	CAEEL
23	39	65.0	321	2	Q61805	CAEBR	Q61805 caenorhabdi	Q61805	CAEBR	Q61805	CAEBR
24	39	65.0	329	2	Q7SHA8	NEUCR	Q7SHA8 neurospora	Q7SHA8	NEUCR	Q7SHA8	NEUCR
25	39	65.0	334	2	Q6N3F8	RHOPA	Q6N3F8 rhodopsendo	Q6N3F8	RHOPA	Q6N3F8	RHOPA
26	39	65.0	537	2	Q59833	SCHPO	Q59833 schizosacch	Q59833	SCHPO	Q59833	SCHPO
27	39	65.0	688	2	Q4HKC7	CAMLA	Q4HKC7 campylobact	Q4HKC7	CAMLA	Q4HKC7	CAMLA
28	39	65.0	759	2	Q88P47	PSEPK	Q88P47 pseudomonas	Q88P47	PSEPK	Q88P47	PSEPK
29	38	63.3	144	2	Q5AEB6	CANAL	Q5AEB6 candida alb	Q5AEB6	CANAL	Q5AEB6	CANAL
30	38	63.3	202	2	Q52BK0	MAGGR	Q52BK0 magnaporthe	Q52BK0	MAGGR	Q52BK0	MAGGR
31	38	63.3	273	2	Q4MTY4	BACCE	Q4MTY4 bacillus ce	Q4MTY4	BACCE	Q4MTY4	BACCE

32	38	63.3	278	2	Q978G1	THEVO	Q978G1 thermoplasma	Q978G1	THEVO	Q978G1	thermoplasma
33	38	63.3	342	2	Q55HF8	CRYNE	Q55HF8 cryptococcus	Q55HF8	CRYNE	Q55HF8	cryptococcus
34	38	63.3	342	2	Q5K6Z7	CRYNE	Q5K6Z7 cryptococcus	Q5K6Z7	CRYNE	Q5K6Z7	cryptococcus
35	38	63.3	345	2	P93588	SOLITU	P93588 solanum tub	P93588	SOLITU	P93588	solanum tub
36	38	63.3	355	2	Q64VT2	BACFR	Q64VT2 bacteroides	Q64VT2	BACFR	Q64VT2	bacteroides
37	38	63.3	357	2	Q5LBC9	BACFN	Q5LBC9 bacteroides	Q5LBC9	BACFN	Q5LBC9	bacteroides
38	38	63.3	457	2	Q7G604	ORYSA	Q7G604 oryza sativ	Q7G604	ORYSA	Q7G604	oryza sativ
39	38	63.3	457	2	Q94LK1	ORYSA	Q94LK1 oryza sativ	Q94LK1	ORYSA	Q94LK1	oryza sativ
40	38	63.3	646	2	Q55RH1	CRYNE	Q55RH1 cryptococcus	Q55RH1	CRYNE	Q55RH1	cryptococcus
41	38	63.3	646	2	Q5KEP9	CRYNE	Q5KEP9 cryptococcus	Q5KEP9	CRYNE	Q5KEP9	cryptococcus
42	38	63.3	660	1	DNLJ	CHLMU	DNLJ chlamydia m	DNLJ	CHLMU	DNLJ	chlamydia m
43	38	63.3	660	2	Q5LSQ2	CHLAB	Q5LSQ2 chlamydia p	Q5LSQ2	CHLAB	Q5LSQ2	chlamydia p
44	38	63.3	662	1	DNLJ	CHLPN	DNLJ chlamydia p	DNLJ	CHLPN	DNLJ	chlamydia p
45	38	63.3	662	2	Q822R2	CHLCV	Q822R2 chlamydia p	Q822R2	CHLCV	Q822R2	chlamydia p

ALIGNMENTS

RESULT 1  
Q9ZRC0 LYCES  
ID Q9ZRC0 LYCES PRELIMINARY; PRT; 320 AA.  
AC Q9ZRC0;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Knotted 3 protein.  
GN Name=TKN3;  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=sp+;  
RA Parnis A., Lifschitz E.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (by similarity).  
DR EMBL; U76408; AAD00252.1; -; mRNA.  
DR HSSP; P41778; ILFU.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR005539; ELK.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR012287; Homeodomain-rel.  
DR InterPro; IPR005540; KNOX1.  
DR InterPro; IPR005541; KNOX2.  
DR Pfam; PF03789; ELK; 1.  
DR Pfam; PF03791; KNOX2; 1.  
DR Pfam; PF03790; KNOX1; 1.  
DR ProDom; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; UNKNOWN\_1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
SQ SEQUENCE 320 AA; 36751 MW; 23C0FB3EEF5C5B5 CRC64;

Query Match 75.0%; Score 45; DB 2; Length 320;  
Best Local Similarity 75.0%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 YDHFRAWF 8  
Db 257 YDHFRAWF 264

RESULT 2  
Q7RNH7 PLAYO  
ID Q7RNH7 PLAYO PRELIMINARY; PRT; 30 AA.  
AC Q7RNH7;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:11:22 ; Search time 183 seconds  
(without alignments)  
24.010 Million cell updates/sec

Title: US-10-602-394A-3

Perfect score: 53

Sequence: 1 YDRFFNAFY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_21:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*
- 9: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	98.1	10	9	ADW04628 Human AGR
2	52	98.1	17	9	ADW04636 Human AGR
3	45	84.9	32	3	AAB00091 Agouti re
4	45	84.9	32	3	AAB00085 Agouti re
5	42	79.2	10	9	ADW04626 Human ago
6	42	79.2	20	3	AAB00104 Agouti re
7	42	79.2	22	3	AAB00100 Agouti re
8	42	79.2	22	3	AAB00092 Agouti re
9	42	79.2	22	3	AAB00099 Agouti re
10	42	79.2	32	3	AAB00082 Agouti re
11	42	79.2	32	3	AAB00089 Agouti re
12	42	79.2	32	3	AAB00083 Agouti re
13	42	79.2	32	3	AAB00084 Agouti re
14	42	79.2	32	3	AAB00088 Agouti re
15	42	79.2	32	3	AAB00090 Agouti re
16	42	79.2	33	4	AAB75127 Human min
17	42	79.2	33	5	AAU74943 Human min
18	42	79.2	34	5	AAU74944 Human min
19	42	79.2	34	5	AAU74945 Human min
20	42	79.2	34	5	AAU75750 Human min
21	42	79.2	46	2	AA49101 Human min
22	42	79.2	46	2	AA49103 Mouse min
23	42	79.2	46	4	AAB75126 Human min
24	42	79.2	46	5	AAU74942 Human min

25	42	79.2	46	9	ADW04672 Human ago
26	42	79.2	47	8	ADP18425 Neurogene
27	42	79.2	48	3	AAB00081 Agouti re
28	42	79.2	50	2	AA33951 Melanocor
29	42	79.2	54	2	AAW26778 Human ago
30	42	79.2	57	2	AAW21862 Human ago
31	42	79.2	57	2	AAW21864 Mouse ago
32	42	79.2	83	2	AAW21863 Mouse ago
33	42	79.2	83	2	AAW21861 Human ago
34	42	79.2	111	2	AAW28598 Chimeric
35	42	79.2	113	2	AAW21866 Amino aci
36	42	79.2	117	2	AAW21867 Amino aci
37	42	79.2	118	2	AAW21856 Amino aci
38	42	79.2	118	2	AAW21865 Amino aci
39	42	79.2	120	2	AAW21857 Amino aci
40	42	79.2	121	2	AAW21858 Amino aci
41	42	79.2	131	2	AAW26779 Mouse ago
42	42	79.2	131	2	AAW49104 Mouse ago
43	42	79.2	131	8	ADS87208 Murine Ag
44	42	79.2	132	2	AAW26777 Human ago
45	42	79.2	132	2	AAW26780 Human ago

#### ALIGNMENTS

RESULT 1  
ADW04628

ID ADW04628 standard; peptide; 10 AA.

XX AC ADW04628;

XX DT 10-MAR-2005 (first entry)

XX DE Human AGRP/MCR agonist chimeric cyclic peptide #2.

XX KW protein engineering; melanocortin receptor; AGRP; agouti related protein;  
XX KW obesity; gene therapy; anorectic; cyclic; melanocortin agonist.

XX OS Homo sapiens.

XX OS Chimeric.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "this residue forms a lactam bridge with Dpr at position 9"

FT Misc-difference 9 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a lactam bridge with Asp at position 2"

FT US2004260063-A1.

XX 23-DEC-2004.

XX 23-JUN-2003; 2003US-00602394.

XX 23-JUN-2003; 2003US-00602394.

XX (HASK/) HASKELL-LUEVANO C.

XX Haskell-Luevano C;

XX WPI; 2005-030733/03.

XX New chimeric peptides and templates based upon melanocortin agonist  
PT peptides and agouti related protein antagonist peptide, useful for  
PT treating or preventing conditions modulated by melanocortin receptors,  
PT such as obesity.

XX Example 3; SEQ ID NO 3; 15pp; English.

XX The present invention relates to a chimeric peptide that is biologically

CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist chimeric cyclic peptide.  
 XX  
 SQ Sequence 10 AA;

Query Match 98.1%; Score 52; DB 9; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDRFFNAFY 10  
 |||||  
 Db 1 YDRFFNAFY 10

RESULT 2  
 ADW04636  
 ID ADW04636 standard; peptide; 17 AA.  
 XX  
 AC ADW04636;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX  
 DE Human AGRP/MCR agonist/NDP-MSH chimeric cyclic peptide.

XX  
 XX protein engineering; melanocortin receptor; AGRP; agouti related protein;  
 XX obesity; gene therapy; anorectic; cyclic; NDP-MSH; melanocortin agonist.  
 XX  
 OS Homo sapiens.  
 OS Chimeric.  
 OS Unidentified.

XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 4 /label= Nle  
 FT Modified-site 4 /note= "Norleucine"  
 FT Misc-difference 6 /note= "This residue forms a lactam bridge with Dpr at  
 FT position 13"  
 FT Misc-difference 13 /note= "2,3-Diaminopropionic acid (Dpr) residue forms a  
 FT lactam bridge with Asp at position 6"  
 FT Modified-site 17 /note= "C-terminal amide"  
 FT  
 XX US2004260063-A1.  
 FN  
 XX 23-DEC-2004.  
 PD  
 XX 23-JUN-2003; 2003US-00602394.  
 PF  
 XX 23-JUN-2003; 2003US-00602394.  
 PR  
 XX (HASK/) HASKELL-LUEVANO C.  
 PA  
 XX Haskell-Luevano C;  
 PI  
 XX WPI; 2005-030733/03.  
 DR  
 XX New chimeric peptides and templates based upon melanocortin agonist  
 XX peptides and agouti related protein antagonist peptide, useful for  
 XX treating or preventing conditions modulated by melanocortin receptors,  
 XX such as obesity.

XX Claim 7; SEQ ID NO 11; 15pp; English.  
 PS  
 XX The present invention relates to a chimeric peptide that is biologically

CC active at melanocortin receptors (MCR) comprising an agouti (ASP) related  
 CC protein (AGRP) template and melanocortin agonist-based bioactive  
 CC determinant sequences which have been substituted for the analogous  
 CC template sequences. The invention is useful for treating or preventing  
 CC various diseases and conditions modulated by melanocortin receptors, such  
 CC as obesity and in gene therapy. The present sequence is the human  
 CC AGRP/MCR agonist/NDP-MSH chimeric cyclic peptide.  
 XX  
 SQ Sequence 17 AA;

Query Match 98.1%; Score 52; DB 9; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDRFFNAFY 10  
 |||||  
 Db 5 YDRFFNAFY 14

RESULT 3  
 AAB00091  
 ID AAB00091 standard; peptide; 32 AA.  
 XX  
 AC AAB00091;  
 XX  
 DT 08-NOV-2000 (first entry)  
 XX  
 DE Agouti related peptide.

XX  
 XX Agouti signalling protein; agouti related peptide; AGRP; ASP; obesity;  
 XX eating disorder; antibody; probe; melanocortin; receptor.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Modified-site 2 /label= Abu  
 FT Modified-site 8 /label= Abu  
 FT Modified-site 10 /label= Abu  
 FT Modified-site 17 /label= Abu  
 FT Modified-site 17 /label= Abu  
 XX WO200044898-A2.  
 PN  
 XX 03-AUG-2000.  
 PD  
 XX 27-JAN-2000; 2000WO-US001879.  
 PF  
 XX 29-JAN-1999; 99US-00240078.  
 PR  
 XX (AMGE-) AMGEN INC.  
 PA  
 XX Jarosinski MA;  
 PI  
 XX WPI; 2000-476223/41.  
 DR  
 XX N-PSDB; AAA47812.  
 DR  
 XX Novel peptides derived from human anti-agouti-related polypeptide and  
 XX anti-agouti-signaling protein, useful for treating eating disorders, e.g.  
 XX obesity.

XX Claim 1; Page 50; 81pp; English.  
 PS  
 XX Agouti related peptides, their fragments, variants or derivatives may be  
 XX used as modulators of feeding behavior (e.g. to treat obesity) and to  
 XX generate anti-agouti-related polypeptide (AGRP) and anti-agouti-signaling  
 XX protein (ASP) antibodies. Nucleic acids encoding these peptides may be  
 XX used as hybridization probes in assays. The antibodies may be used to  
 XX inhibit the binding of the AGRP/ASP peptide to melanocortin receptors or  
 XX in vitro or in vivo diagnostics. This sequence corresponds to amino  
 XX acids 81-112 of human AGRP with amino acids 82, 88, 90 and 97 of native



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:18:47 ; Search time 47 Seconds  
(without alignments)  
17.591 Million cell updates/sec

Title: US-10-602-394A-3

Perfect score: 53

Sequence: 1 YDRFFNAFYX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/PCUS\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	84.9	32	2	US-09-240-078-5
2	45	84.9	32	2	US-09-240-078-11
3	42	79.2	20	2	US-09-240-078-24
4	42	79.2	22	2	US-09-240-078-12
5	42	79.2	22	2	US-09-240-078-19
6	42	79.2	22	2	US-09-240-078-20
7	42	79.2	32	2	US-09-240-078-2
8	42	79.2	32	2	US-09-240-078-3
9	42	79.2	32	2	US-09-240-078-4
10	42	79.2	32	2	US-09-240-078-8
11	42	79.2	32	2	US-09-240-078-9
12	42	79.2	32	2	US-09-240-078-10
13	42	79.2	48	2	US-09-240-078-1
14	42	79.2	50	2	US-09-031-902-2
15	42	79.2	54	1	US-08-757-541-8
16	42	79.2	54	2	US-09-033-275-8
17	42	79.2	54	2	US-09-342-581-8
18	42	79.2	57	2	US-09-581-894A-7
19	42	79.2	57	2	US-09-581-894A-9
20	42	79.2	83	2	US-09-581-894A-6
21	42	79.2	83	2	US-09-581-894A-8
22	42	79.2	113	2	US-09-581-894A-11
23	42	79.2	117	2	US-09-581-894A-12
24	42	79.2	118	2	US-09-581-894A-1
25	42	79.2	118	2	US-09-581-894A-10
26	42	79.2	120	2	US-09-581-894A-2
27	42	79.2	121	2	US-09-581-894A-3

28	79.2	131	1	US-08-757-541-10	Sequence 10, Appl
29	79.2	131	2	US-09-033-275-10	Sequence 10, Appl
30	79.2	131	2	US-09-342-581-10	Sequence 10, Appl
31	79.2	132	1	US-08-757-541-7	Sequence 7, Appl
32	79.2	132	1	US-08-757-541-11	Sequence 11, Appl
33	79.2	132	2	US-09-033-275-7	Sequence 7, Appl
34	79.2	132	2	US-09-033-275-11	Sequence 11, Appl
35	79.2	132	2	US-09-342-581-7	Sequence 7, Appl
36	79.2	132	2	US-09-342-581-11	Sequence 11, Appl
37	79.2	132	2	US-09-384-302A-10	Sequence 10, Appl
38	79.2	605	2	US-09-581-894A-15	Sequence 15, Appl
39	79.2	620	2	US-09-581-894A-13	Sequence 13, Appl
40	79.2	621	2	US-09-581-894A-17	Sequence 17, Appl
41	79.2	621	2	US-09-581-894A-19	Sequence 19, Appl
42	79.2	654	2	US-09-581-894A-4	Sequence 4, Appl
43	79.2	666	2	US-09-581-894A-16	Sequence 16, Appl
44	79.2	683	2	US-09-581-894A-14	Sequence 14, Appl
45	79.2	684	2	US-09-581-894A-18	Sequence 18, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-240-078-5

; Sequence 5, Application US/09240078

; Patent No. 6303749

; GENERAL INFORMATION:

; APPLICANT: Jarosinski, Mark A.

; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs

; FILE REFERENCE: A-569

; CURRENT APPLICATION NUMBER: US/09/240,078

; CURRENT FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 32

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide

; OTHER INFORMATION: Analog with synthetic amino acid (amino butyric

; OTHER INFORMATION: acid) at positions 12, 15, 20 and 27 and labeled

; OTHER INFORMATION: as Xaa.

US-09-240-078-5

Query Match

Best Local Similarity 84.9%; Score 45; DB 2; Length 32;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDRFFNAFYX 10

Db 19 YXRFNFAFYX 28

##### RESULT 2

US-09-240-078-11

; Sequence 11, Application US/09240078

; Patent No. 6303749

; GENERAL INFORMATION:

; APPLICANT: Jarosinski, Mark A.

; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs

; FILE REFERENCE: A-569

; CURRENT APPLICATION NUMBER: US/09/240,078

; CURRENT FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 32

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide

```
; OTHER INFORMATION: Analog with synthetic amino acid (amino butyric
; OTHER INFORMATION: acid) located at positions 2, 8, 10 and 17 and
; OTHER INFORMATION: labeled as Xaa.
US-09-240-078-11

Query Match      84.9%; Score 45; DB 2; Length 32;
Best Local Similarity 90.0%; Pred. No. 0.22; 1; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 YDRFFNAFY 10
Db 9 YCRFFNAFY 18
    |||||

RESULT 3
US-09-240-078-24
; Sequence 24, Application US/09240078
; Patent No. 6303749
; GENERAL INFORMATION:
; APPLICANT: Jarosinski, Mark A.
; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs
; FILE REFERENCE: A-569
; CURRENT APPLICATION NUMBER: US/09/240,078
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide
; OTHER INFORMATION: Analog with amino acids located at positions 3, 6,
; OTHER INFORMATION: and 17 replaced by alanine (Ala) and in which the
; OTHER INFORMATION: amino terminus is acetylated..
US-09-240-078-24

Query Match      79.2%; Score 42; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 0.47; 2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy 1 YDRFFNAFY 10
Db 7 YCRFFNAFY 16
    |||||

RESULT 4
US-09-240-078-12
; Sequence 12, Application US/09240078
; Patent No. 6303749
; GENERAL INFORMATION:
; APPLICANT: Jarosinski, Mark A.
; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs
; FILE REFERENCE: A-569
; CURRENT APPLICATION NUMBER: US/09/240,078
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide
; OTHER INFORMATION: Analog with synthetic amino acid (amino butyric
; OTHER INFORMATION: acid) located at positions 2, 5, and 8 and labeled
; OTHER INFORMATION: as Xaa.
US-09-240-078-12

Query Match      79.2%; Score 42; DB 2; Length 22;
Best Local Similarity 80.0%; Pred. No. 0.52; 2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy 1 YDRFFNAFY 10
Db 9 YCRFFNAFY 18
    |||||

US-09-240-078-11
```

```
Qy 1 YDRFFNAFY 10
Db 9 YCRFFNAFY 18
    |||||

RESULT 5
US-09-240-078-19
; Sequence 19, Application US/09240078
; Patent No. 6303749
; GENERAL INFORMATION:
; APPLICANT: Jarosinski, Mark A.
; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs
; FILE REFERENCE: A-569
; CURRENT APPLICATION NUMBER: US/09/240,078
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide
; OTHER INFORMATION: Analog with amino acids located at positions 2, 5
; OTHER INFORMATION: and 8 replaced by alanine (Ala).
US-09-240-078-19

Query Match      79.2%; Score 42; DB 2; Length 22;
Best Local Similarity 80.0%; Pred. No. 0.52; 2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy 1 YDRFFNAFY 10
Db 9 YCRFFNAFY 18
    |||||

RESULT 6
US-09-240-078-20
; Sequence 20, Application US/09240078
; Patent No. 6303749
; GENERAL INFORMATION:
; APPLICANT: Jarosinski, Mark A.
; TITLE OF INVENTION: No. 6303749el Agouti and Agouti-Related Peptide Analogs
; FILE REFERENCE: A-569
; CURRENT APPLICATION NUMBER: US/09/240,078
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AGRP Peptide
; OTHER INFORMATION: Analog with amino acids located at positions 2, 3,
; OTHER INFORMATION: 5 and 8 replaced by alanine (Ala).
US-09-240-078-20

Query Match      79.2%; Score 42; DB 2; Length 22;
Best Local Similarity 80.0%; Pred. No. 0.52; 2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy 1 YDRFFNAFY 10
Db 9 YCRFFNAFY 18
    |||||

RESULT 7
US-09-240-078-2
; Sequence 2, Application US/09240078
; Patent No. 6303749
; GENERAL INFORMATION:
; APPLICANT: Jarosinski, Mark A.
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:29:57 ; Search time 164 Seconds  
(without alignments)  
25.477 Million cell updates/sec

Title: US-10-602-394A-3  
Perfect score: 53  
Sequence: 1 YDRFFNAFY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569.

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	33	3	US-09-851-586A-3
2	42	79.2	34	3	US-09-851-586A-4
3	42	79.2	34	3	US-09-851-586A-5
4	42	79.2	46	3	US-09-851-586A-2
5	42	79.2	47	5	US-10-718-071-39
6	42	79.2	54	3	US-09-754-862-8
7	42	79.2	54	4	US-10-256-590-8
8	42	79.2	131	3	US-09-754-862-10
9	42	79.2	131	4	US-10-256-590-10
10	42	79.2	132	3	US-09-754-862-7
11	42	79.2	132	3	US-09-754-862-11
12	42	79.2	132	3	US-09-851-586A-1
13	42	79.2	132	4	US-10-207-330-10
14	42	79.2	132	4	US-10-256-590-7
15	42	79.2	132	4	US-10-256-590-11
16	42	79.2	132	4	US-10-379-747-26
17	42	79.2	132	6	US-11-060-291-10
18	40	75.5	484	4	US-10-425-114-49505
19	40	75.5	728	4	US-10-424-599-189465
20	39	73.6	339	5	US-10-739-930-5959
21	37	69.8	604	4	US-10-424-599-207219
22	37	69.8	1548	4	US-10-437-963-112738
23	36	67.9	38	4	US-10-424-599-200686
24	36	67.9	360	4	US-10-282-122A-68814
25	36	67.9	460	4	US-10-369-493-5817
26	36	67.9	733	4	US-10-437-963-171240
27	36	67.9	812	4	US-10-437-963-123561

28	36	67.9	885	4	US-10-437-963-171263	Sequence 171263,
29	35	66.0	117	3	US-09-881-752A-232	Sequence 232, Appl
30	35	66.0	516	3	US-09-430-029-5	Sequence 5, Appl
31	35	66.0	800	6	US-11-097-143-15480	Sequence 15480, A
32	35	66.0	1239	4	US-10-425-115-303754	Sequence 303754, A
33	35	66.0	1683	4	US-10-369-493-22273	Sequence 22273, A
34	34	64.2	12	5	US-10-771-232-19	Sequence 19, Appl
35	34	64.2	62	4	US-10-424-599-245398	Sequence 245398,
36	34	64.2	71	4	US-10-437-963-200519	Sequence 200519,
37	34	64.2	82	4	US-10-425-115-34841	Sequence 34841,
38	34	64.2	101	4	US-10-282-122A-57526	Sequence 57526, A
39	34	64.2	138	4	US-10-767-701-43254	Sequence 43254, A
40	34	64.2	155	4	US-10-424-599-278805	Sequence 278805,
41	34	64.2	201	4	US-10-282-122A-67733	Sequence 67733, A
42	34	64.2	202	4	US-10-282-122A-48979	Sequence 48979, A
43	34	64.2	207	4	US-10-282-122A-69731	Sequence 69731, A
44	34	64.2	208	4	US-10-282-122A-47739	Sequence 47739, A
45	34	64.2	208	4	US-10-282-122A-54542	Sequence 54542, A

ALIGNMENTS

RESULT 1  
US-09-851-586A-3  
; Sequence 3, Application US/09851586A  
; Publication No. US20030064921A1  
; GENERAL INFORMATION:  
; APPLICANT: MILLHAUSER, GLENN  
; APPLICANT: THOMPSON, DARREN  
; APPLICANT: BOLIN, KIMBERLEY  
; APPLICANT: ANDERSON, JOE  
; APPLICANT: MCNUITY, JOSEPH  
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR MODULATING MELANOCORTIN RECEPTOR LIGAN

; FILE REFERENCE: 407t-980910US  
; CURRENT APPLICATION NUMBER: US/09/851,586A  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: PCT/US99/25201  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/203,271  
; PRIOR FILING DATE: 2000-05-09  
; PRIOR APPLICATION NUMBER: 60/226,047  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-851-586A-3

Query Match 79.2%; Score 42; DB 3; Length 33;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YDRFFNAFY 10  
Db 23 YCRFFNAFCY 32

RESULT 2  
US-09-851-586A-4  
; Sequence 4, Application US/09851586A  
; Publication No. US20030064921A1  
; GENERAL INFORMATION:  
; APPLICANT: MILLHAUSER, GLENN  
; APPLICANT: THOMPSON, DARREN  
; APPLICANT: BOLIN, KIMBERLEY  
; APPLICANT: ANDERSON, JOE  
; APPLICANT: MCNUITY, JOSEPH  
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR MODULATING MELANOCORTIN RECEPTOR LIGAN

FILE REFERENCE: 407t-980910US  
CURRENT APPLICATION NUMBER: US/09/851,586A  
CURRENT FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: PCT/US99/25201  
PRIOR FILING DATE: 1999-10-27  
PRIOR APPLICATION NUMBER: 60/203,271  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: 60/226,047  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic mini-AGRP  
US-09-851-586A-4

Query Match 79.2%; Score 42; DB 3; Length 34;  
Best Local Similarity 80.0%; Pred. No. 1.5; Length 34;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDRFFNAFY 10  
| | | | | | | |  
Db 23 YCRFFNAFCY 32

RESULT 3  
US-09-851-586A-5  
Sequence 5, Application US/09851586A  
Publication No. US20030064921A1  
GENERAL INFORMATION:  
APPLICANT: MILLHAUSER, GLENN  
APPLICANT: THOMPSON, DARREN  
APPLICANT: BOLIN, KIMBERLEY  
APPLICANT: ANDERSON, JOE  
APPLICANT: MCNUITY, JOSEPH  
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR MODULATING MELANOCORTIN RECEPTOR LIGAND  
TITLE OF INVENTION: AND ACTIVITY  
FILE REFERENCE: 407t-980910US  
CURRENT APPLICATION NUMBER: US/09/851,586A  
CURRENT FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: PCT/US99/25201  
PRIOR FILING DATE: 1999-10-27  
PRIOR APPLICATION NUMBER: 60/203,271  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: 60/226,047  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic AGRP fragment  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: C blocked with acetyl  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (34)..(34)  
OTHER INFORMATION: R blocked with amino  
US-09-851-586A-5

Query Match 79.2%; Score 42; DB 3; Length 34;  
Best Local Similarity 80.0%; Pred. No. 1.5; Length 34;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDRFFNAFY 10  
| | | | | | | |

Db 23 YCRFFNAFCY 32

RESULT 4

US-09-851-586A-2  
Sequence 2, Application US/09851586A  
Publication No. US20030064921A1  
GENERAL INFORMATION:  
APPLICANT: MILLHAUSER, GLENN  
APPLICANT: THOMPSON, DARREN  
APPLICANT: BOLIN, KIMBERLEY  
APPLICANT: ANDERSON, JOE  
APPLICANT: MCNUITY, JOSEPH  
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR MODULATING MELANOCORTIN RECEPTOR LIGAND  
TITLE OF INVENTION: AND ACTIVITY  
FILE REFERENCE: 407t-980910US  
CURRENT APPLICATION NUMBER: US/09/851,586A  
CURRENT FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: PCT/US99/25201  
PRIOR FILING DATE: 1999-10-27  
PRIOR APPLICATION NUMBER: 60/203,271  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: 60/226,047  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-851-586A-2

Query Match 79.2%; Score 42; DB 3; Length 46;  
Best Local Similarity 80.0%; Pred. No. 2; Length 46;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDRFFNAFY 10  
| | | | | | | |  
Db 23 YCRFFNAFCY 32

RESULT 5  
US-10-718-071-39  
Sequence 39, Application US/10718071  
Publication No. US20050009847A1  
GENERAL INFORMATION:  
APPLICANT: Bertilsson, Goran  
APPLICANT: Erlandsson, Rikard  
APPLICANT: Frisen, Jonas  
APPLICANT: Haegerstrand, Anders  
APPLICANT: Heidrich, Jessica  
APPLICANT: Hellstrom, Kristina  
APPLICANT: Haggblad, Johan  
APPLICANT: Jansson, Katarina  
APPLICANT: Kortesmaa, Jarkko  
APPLICANT: Lindquist, Per  
APPLICANT: Lundh, Hanna  
APPLICANT: McGuire, Jacqueline  
APPLICANT: Mercer, Alex  
APPLICANT: Nyberg, Karl  
APPLICANT: Ossolinak, Amina  
APPLICANT: Patrone, Cesare  
APPLICANT: Ronnhelm, Harriet  
APPLICANT: Wikstrom, Lillian  
APPLICANT: Zachrisson, Olof  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INCREASING NEUROGENESIS  
FILE REFERENCE: 21882-517 UTIL  
CURRENT APPLICATION NUMBER: US/10/718,071  
CURRENT FILING DATE: 2003-11-20  
PRIOR APPLICATION NUMBER: US 60/427,912  
PRIOR FILING DATE: 2002-11-20  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: PatentIn version 3.2

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:30:32 ; Search time 16 Seconds  
(without alignments)  
8.202 Million cell updates/sec

Title: US-10-602-394A-3

Perfect score: 53

Sequence: 1 YDRFFNAFY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	46	7	US-11-188-552-70
2	42	79.2	50	7	US-11-188-552-69
3	42	79.2	50	7	US-11-174-845-2
4	42	79.2	50	7	US-11-174-851-2
5	42	79.2	131	6	US-10-518-955-2
6	42	79.2	132	6	US-10-518-955-4
7	34	64.2	305	7	US-11-156-084-186
8	34	64.2	389	7	US-11-129-143-81
9	33	62.3	32	6	US-10-895-064-2411
10	33	62.3	109	7	US-11-156-084-139
11	33	62.3	286	7	US-11-057-012-93
12	33	62.3	350	6	US-10-467-657-3588
13	32	60.4	267	7	US-11-052-554A-300
14	31	58.5	233	6	US-10-467-657-8268
15	31	58.5	286	6	US-10-793-626-2192
16	31	58.5	691	6	US-10-467-657-7170
17	30.5	57.5	607	6	US-10-131-826A-344
18	30.5	57.5	607	7	US-11-176-863-2
19	30	56.6	350	6	US-10-467-657-224
20	30	56.6	350	6	US-10-467-657-512
21	30	56.6	375	6	US-10-517-939-2
22	30	56.6	802	6	US-10-510-386-2
23	30	56.6	1400	6	US-10-821-234-1045
24	29	54.7	134	6	US-10-793-626-3006
25	29	54.7	134	6	US-10-793-626-3046

```

26 29 54.7 134 6 US-10-793-626-3304 Sequence 3304, Ap
27 29 54.7 162 5 US-09-978-360A-732 Sequence 732, App
28 29 54.7 267 6 US-10-873-528-101 Sequence 101, App
29 29 54.7 321 6 US-10-467-657-2352 Sequence 2352, Ap
30 29 54.7 322 7 US-11-156-084-318 Sequence 318, App
31 29 54.7 354 6 US-10-467-657-4352 Sequence 4352, Ap
32 29 54.7 360 7 US-11-194-246-321 Sequence 321, App
33 29 54.7 424 7 US-11-055-822-1036 Sequence 1036, Ap
34 29 54.7 497 6 US-10-454-437-410 Sequence 410, App
35 29 54.7 644 6 US-10-467-657-5958 Sequence 5958, Ap
36 29 54.7 757 7 US-11-097-463-2 Sequence 2, Appli
37 29 54.7 844 7 US-11-097-463-4 Sequence 4, Appli
38 29 54.7 846 6 US-10-517-939-90 Sequence 90, Appl
39 29 54.7 1102 7 US-11-098-686-10951 Sequence 10951, A
40 28 52.8 114 5 US-09-978-360A-726 Sequence 726, App
41 28 52.8 116 6 US-10-467-657-7486 Sequence 7486, Ap
42 28 52.8 118 7 US-11-049-536-606 Sequence 606, App
43 28 52.8 138 7 US-11-098-686-11366 Sequence 11366, A
44 28 52.8 147 6 US-10-485-517-343 Sequence 343, App
45 28 52.8 155 6 US-10-467-657-1206 Sequence 1206, Ap

```

#### ALIGNMENTS

```

RESULT 1
US-11-188-552-70
; Sequence 70, Application US/11188552
; Publication No. US20050282739A1
; GENERAL INFORMATION:
; APPLICANT: Palatin Technologies, Inc.
; APPLICANT: Sharma, Shubh
; APPLICANT: Shi, Yi-Qun
; APPLICANT: Bastos, Margarita
; APPLICANT: Rajpurohit, Ramesh
; APPLICANT: Cai, Hui-Zhi
; TITLE OF INVENTION: Metallopeptide Compounds
; FILE REFERENCE: 0507-05
; CURRENT APPLICATION NUMBER: US/11/188,552
; CURRENT FILING DATE: 2005-07-25
; PRIOR APPLICATION NUMBER: US 60/590,933
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/769,695
; PRIOR FILING DATE: 2005-01-30
; PRIOR APPLICATION NUMBER: US 60/444,129
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/464,117
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/50075
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/256,842
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/304,835
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 60/327,835
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-188-552-70

```

Query Match 79.2%; Score 42; DB 7; Length 46;  
Best Local Similarity 80.0%; Pred. No. 0.15;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 YDRFFNAFY 10
Db 23 YCRFFNAFCY 32

```

```
RESULT 2
US-11-188-552-69
; Sequence 69, Application US/11188552
; Publication No. US20050282739A1
; GENERAL INFORMATION:
; APPLICANT: Palatin Technologies, Inc.
; APPLICANT: Sharma, Shubh
; APPLICANT: Shi, Yi-Qun
; APPLICANT: Bascos, Margarita
; APPLICANT: Rajpurohit, Ramesh
; APPLICANT: Cai, Hui-Zhi
; TITLE OF INVENTION: Metallopeptide Compounds
; FILE REFERENCE: 0507-05
; CURRENT APPLICATION NUMBER: US/11/188,552
; CURRENT FILING DATE: 2005-07-25
; PRIOR APPLICATION NUMBER: US 60/590,933
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/769,695
; PRIOR FILING DATE: 2005-01-30
; PRIOR APPLICATION NUMBER: US 60/444,129
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/464,117
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/50075
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/256,842
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/304,835
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 60/327,835
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-188-552-69

Query Match      79.2%; Score 42; DB 7; Length 50;
Best Local Similarity 80.0%; Pred. No. 0.16;
Matches      8; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

Qy      1 YDRFFNAFY 10
      |||||||
Db      27 YCRFFNAFCY 36

RESULT 3
US-11-174-845-2
; Sequence 2, Application US/11174845
; Publication No. US20060014676A1
; GENERAL INFORMATION:
; APPLICANT: Palatin Technologies, Inc.
; APPLICANT: Sharma, Shubh D.
; APPLICANT: Rajpurohit, Ramesh
; APPLICANT: Shaddack, Annette M.
; APPLICANT: Shi, Yi-Qun
; APPLICANT: Burtis, Kevin B.
; TITLE OF INVENTION: Cyclic Peptides for Treatment of Cachexia
; FILE REFERENCE: 70025-US05-071
; CURRENT APPLICATION NUMBER: US/11/174,845
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: US 60/585,791
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US 10/638,071
; PRIOR FILING DATE: 2003-08-03
; PRIOR APPLICATION NUMBER: PCT/US02/22196
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-845-2

Query Match      79.2%; Score 42; DB 7; Length 50;
Best Local Similarity 80.0%; Pred. No. 0.16;
Matches      8; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

Qy      1 YDRFFNAFY 10
      |||||||
Db      27 YCRFFNAFCY 36

RESULT 4
US-11-174-851-2
; Sequence 2, Application US/11174851
; Publication No. US20060014194A1
; GENERAL INFORMATION:
; APPLICANT: Palatin Technologies, Inc.
; APPLICANT: Sharma, Shubh D.
; APPLICANT: Rajpurohit, Ramesh
; APPLICANT: Shaddack, Annette M.
; APPLICANT: Shi, Yi-Qun
; APPLICANT: Burtis, Kevin B.
; TITLE OF INVENTION: Cyclic Peptides for Treatment of Cachexia
; FILE REFERENCE: 70025-US05-0707
; CURRENT APPLICATION NUMBER: US/11/174,851
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: US 60/585,791
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US 10/638,071
; PRIOR FILING DATE: 2003-08-03
; PRIOR APPLICATION NUMBER: PCT/US02/22196
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-851-2

Query Match      79.2%; Score 42; DB 7; Length 50;
Best Local Similarity 80.0%; Pred. No. 0.16;
Matches      8; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

Qy      1 YDRFFNAFY 10
      |||||||
Db      27 YCRFFNAFCY 36

RESULT 5
US-10-518-955-2
; Sequence 2, Application US/10518955
; Publication No. US20050257279A1
; GENERAL INFORMATION:
; APPLICANT: Qian, Su
; APPLICANT: Van der Ploeg, Leonardus, H.T.
; APPLICANT: Chen, Howard
; APPLICANT: Weingarth, Drew T.
; APPLICANT: Trumbauer, Myrna
; APPLICANT: Metzger, Joseph M.
; TITLE OF INVENTION: Agouti-related protein deficient cells,
; TITLE OF INVENTION: non-human transgenic animals and methods of selecting
; TITLE OF INVENTION: compounds which regulate energy metabolism
; FILE REFERENCE: 21033YP
; CURRENT APPLICATION NUMBER: US/10/518,955
; CURRENT FILING DATE: 2004-12-17
; PRIOR APPLICATION NUMBER: PCT/US03/20245
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/393,391
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 14
; LENGTH: 50
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:14:52 ; Search time 38 Seconds  
(without alignments)  
25.320 Million cell updates/sec

Title: US-10-602-394A-3  
Perfect score: 53  
Sequence: 1 YDRFFNAFY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	75.5	548	2	T18761	hypothetical prote
2	38	71.7	118	2	D71818	hypothetical prote
3	38	71.7	282	2	F86786	GTP-binding protei
4	37	69.8	653	2	D69815	conserved hypotet
5	36	67.9	389	2	S31123	hypothetical prote
6	36	67.9	408	2	H71672	ampg protein (ampG
7	36	67.9	460	2	F88544	protein F59B2.2 [i
8	36	67.9	3394	2	T18501	hypothetical prote
9	35	66.0	117	2	A84701	conserved hypotet
10	35	66.0	252	2	B71105	hypothetical prote
11	35	66.0	301	2	T24993	hypothetical prote
12	35	66.0	392	2	B70242	conserved hypotet
13	35	66.0	641	2	JC7142	dextranase [EC 3.2
14	35	66.0	863	2	C36964	probable permease,
15	35	66.0	1683	2	S38103	SPO14 protein - ye
16	34	64.2	160	2	S77324	hypothetical prote
17	34	64.2	208	2	B71327	hypothetical prote
18	34	64.2	234	2	AF3543	phosphoglycolate p
19	34	64.2	271	2	H70888	hypothetical prote
20	34	64.2	280	2	F70805	hypothetical prote
21	34	64.2	305	2	C72366	tRNA delta-2-isope
22	34	64.2	328	2	AI3541	NADH2 dehydrogenas
23	34	64.2	478	2	H50246	hypothetical prote
24	34	64.2	499	2	S52422	chitinase [EC 3.2.
25	34	64.2	499	2	S04856	chitinase [EC 3.2.
26	34	64.2	616	2	S64156	probable membrane
27	34	64.2	653	2	AG1190	hypothetical trans
28	34	64.2	653	2	AG1548	hypothetical trans
29	34	64.2	913	2	JG0168	gob-5 protein - mo

30	64.2	1596	2	AG2501	hypothetical prote
31	64.2	4981	2	T18489	hypothetical prote
32	62.3	104	2	S59108	hypothetical prote
33	62.3	168	2	D87310	hypothetical prote
34	62.3	205	2	H86767	hypothetical prote
35	62.3	248	2	A46652	glucosamine-6-phos
36	62.3	281	2	G87451	hypothetical prote
37	62.3	283	2	C98002	conserved hypotet
38	62.3	283	2	G95133	GTP-binding protei
39	62.3	294	2	A57478	serine O-acetyltra
40	62.3	298	2	T33317	hypothetical prote
41	62.3	341	2	I57997	hypothetical calci
42	62.3	343	2	D69482	hypothetical prote
43	62.3	396	2	T33315	hypothetical prote
44	62.3	415	2	D95248	conserved hypotet
45	62.3	415	2	A98113	conserved hypotet

ALIGNMENTS

RESULT 1

T18761  
hypothetical protein B0462.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18761  
R;Mortimore, B.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19018  
A;Accession: T18761  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-548 <WIL>  
A;Cross-references: UNIPROT:O62019; UNIPARC:UPI00000077D6B; EMBL:Z81456; PIDN:CAB03810.1  
C;Experimental source: clone B0462  
C;Genetics:  
A;Gene: CESP:B0462.1  
A;Map position: 5  
A;Introns: 21/1; 93/3; 135/3; 164/3; 289/3; 372/2

Query Match 75.5%; Score 40; DB 2; Length 548;  
Best Local Similarity 77.8%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DRFPNAPXY 10  
Db 539 ERFNAPCY 547

RESULT 2

D71818  
hypothetical protein jhp1342 - Helicobacter pylori (strain J99)  
C;Species: Helicobacter pylori  
A;Variety: strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: D71818  
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Millis, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: D71818  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-118 <ARN>  
A;Cross-references: UNIPROT:Q9ZJG9; UNIPARC:UPI0000013AAE7; GB:AE001557; GB:AE001439; NI  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp1342  
C;Superfamily: conserved hypothetical protein HI1000  
Query Match 71.7%; Score 38; DB 2; Length 118;





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2006, 21:11:37 ; Search time 229 Seconds  
(without alignments)  
30.809 Million cell updates/sec

Title: US-10-602-394A-3  
Perfect score: 53  
Sequence: 1 YDRFFNAPFY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	71	2 Q90WY7 COTJA	Q90WY7 coturnix co
2	42	79.2	105	2 Q6J648 SHEEP	Q6J648 ovis aries
3	42	79.2	121	2 Q9GLM5 PIG	Q9GLM5 sus scrofa
4	42	79.2	131	1 AGRP MOUSE	P56473 mus musculus
5	42	79.2	132	1 AGRP HUMAN	O00253 homo sapien
6	42	79.2	134	1 AGRP BOVIN	P56413 bos taurus
7	42	79.2	134	1 AGRP PIG	Q9TUL8 sus scrofa
8	42	79.2	154	2 Q9PWG2 CHICK	Q9PWG2 gallus gall
9	42	79.2	165	2 Q9W7R0 CHICK	Q9W7R0 gallus gall
10	41	77.4	2133	2 Q813V5 PLAF7	Q813V5 plasmodium
11	40	75.5	186	2 Q7NBA0 MYCGA	Q7NBA0 mycoplasma
12	40	75.5	548	2 Q62019 CABEL	O62019 caenorhabdi
13	40	75.5	575	2 Q4P6K3 USTWA	Q4P6K3 ustilago ma
14	39	73.6	339	1 PER29 ARATH	Q91SP0 arabidopsis
15	39	73.6	877	2 Q7RGA6 PLAYO	Q7RGA6 plasmodium
16	39	73.6	940	2 Q4Z453 PLABE	Q4Z453 plasmodium
17	38	71.7	118	1 Y849 HELPJ	Q9ZJ99 helicobacte
18	38	71.7	196	2 Q4HN72 CAMLA	Q4HN72 campylobact
19	38	71.7	282	2 Q9CG16 LACLA	Q9CG16 lactococcus
20	38	71.7	283	2 Q8DZT4 STRA5	Q8DZT4 streptococc
21	38	71.7	283	2 Q8E511 STRA3	Q8E511 streptococc
22	38	71.7	284	2 Q5M022 STRT1	Q5M022 streptococc
23	38	71.7	284	2 Q5M4N7 STRT2	Q5M4N7 streptococc
24	38	71.7	331	2 Q5KD08 CRINE	Q5KD08 cryptococcu
25	38	71.7	333	2 Q55PK8 CRINE	Q55PK8 cryptococcu
26	38	71.7	519	2 Q72M60 LEPIC	Q72M60 leptospira
27	38	71.7	519	2 Q8EYN9 LEPIN	Q8EYN9 leptospira
28	38	71.7	829	2 Q4Z5E3 PLABE	Q4Z5E3 plasmodium
29	38	71.7	953	2 Q4Y0L8 PLACH	Q4Y0L8 plasmodium
30	37	69.8	86	2 Q7YRW0 SHEEP	Q7YRW0 ovis aries
31	37	69.8	104	2 Q5WDR3 BACSK	Q5WDR3 bacillus cl

```

32 37 69.8 174 2 Q7RNI9 PLAYO Q7RNI9 plasmodium
33 37 69.8 338 2 Q626B2 CAEBR Q626B2 caenorhabdi
34 37 69.8 499 2 Q4RVL5 TETNG Q4RVL5 tetradon n
35 37 69.8 520 2 Q9FVY8 ORISA Q9FVY8 oryza sativ
36 37 69.8 590 2 Q57Y41 9TRYP Q57Y41 trypanosoma
37 37 69.8 636 2 Q55FG4 DICDI Q55FG4 dictyosteli
38 37 69.8 639 2 Q797B3 BACSU Q797B3 bacillus su
39 37 69.8 639 2 Q65D92 BACLD Q65D92 bacillus li
40 37 69.8 653 2 Q06487 BACSU Q06487 bacillus su
41 37 69.8 677 2 Q7RE75 PLAYO Q7RE75 plasmodium
42 37 69.8 1437 2 Q7Y98 CRYPV Q7Y98 cryptospori
43 37 69.8 1475 2 Q5CLD0 CRYHO Q5CLD0 cryptospori
44 37 69.8 1476 2 Q5CX72 CRYPV Q5CX72 cryptospori
45 36 67.9 50 2 Q4SEW0 TETNG Q4SEW0 tetradon n

```

#### ALIGNMENTS

##### RESULT 1

```

Q90WY7 COTJA
ID Q90WY7_COTJA PRELIMINARY; PRT; 71 AA.
AC Q90WY7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Agouti-related protein (Fragment).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hypothalamus;
RA Boswell T., Li Q., Takeuchi S.;
RX MEDLINE=22003442; PubMed=12008019; DOI=10.1016/S0169-328X(02)00145-6;
RT "Neurons expressing neuropeptide Y mRNA in the infundibular
RT hypothalamus of Japanese quail are activated by fasting and co-express
RT agouti-related protein mRNA.";
RL Brain Res. Mol. Brain Res. 100:31-42(2002).
DR EMBL; AY048849; AAL06600.1; -; mRNA.
DR HSSP; C00253; 1MR0.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009755; P:hormone-mediated signaling; IEA.
DR InterPro; IPR007733; Agouti.
DR Pfam; PF05039; Agouti; 1.
FT NON TER 1
SQ SEQUENCE 71 AA; 7866 MW; BEF043C9E7A71B22 CRC64;

```

Query Match 79.2%; Score 42; DB 2; Length 71;  
Best Local Similarity 80.0%; Pred. No. 2.5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDRFFNAPFY 10

Db 49 YCRFFNAPFY 58

##### RESULT 2

```

Q6J648 SHEEP
ID Q6J648_SHEEP PRELIMINARY; PRT; 105 AA.
AC Q6J648;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Agouti-related peptide (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RA Wagner C.G., McMahon C.D., Marks D.L.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY596306; AAT41659.1; -; Genomic_DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0009755; P:hormone-mediated signaling; IEA.
DR InterPro; IPR007733; Agouti.
DR Pfam; PF05039; Agouti; 1.
DR NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11665 MW; EP54258B2ED81461 CRC64;

Query Match 79.2%; Score 42; DB 2; Length 105;
Best Local Similarity 80.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QV 1 YDRFFNAFY 10
Db 88 YCRFFNAFY 97

RESULT 3
Q9GLM5_PIG
ID Q9GLM5_PIG PRELIMINARY; PRT; 121 AA.
AC Q9GLM5
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Agouti-related protein (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=HYPOTHALAMUS;
RA Materi R.L., Dyer C.J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF220531; AAG09464.1; -; mRNA.
DR HSSP; O00253; 1HYK.
DR SMR; Q9GLM5; 76-121.
DR GO; GO:0009755; C:extracellular region; IEA.
DR GO; GO:0005576; P:hormone-mediated signaling; IEA.
DR InterPro; IPR007733; Agouti.
DR Pfam; PF05039; Agouti; 1.
FT NON_TER 1
SQ SEQUENCE 121 AA; 13276 MW; DD736F01B5B5766E CRC64;

Query Match 79.2%; Score 42; DB 2; Length 121;
Best Local Similarity 80.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QV 1 YDRFFNAFY 10
Db 98 YCRFFNAFY 107

RESULT 4
ID_AGRP_MOUSE
ID_AGRP_MOUSE STANDARD; PRT; 131 AA.
AC P56473; O35967;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Agouti-related protein precursor.
GN Name=AgRP; Synonyms=AgRP, Art;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129;
RX MEDLINE=97458244; PubMed=9311920; DOI=10.1126/science.278.5335.135;
RA Ollmann M.M., Wilson B.D., Yang Y.K., Kerns J.A., Chen Y., Gantz I.,
RA Barsh G.S.;
RT "Antagonism of central melanocortin receptors in vitro and in vivo by
RT agouti-related protein.";
RL Science 278:135-138(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97230362; PubMed=9119224;
RA Shutter J.R., Graham M., Kinsey A.C., Scully S., Luethy R.,
RA Stark K.L.;
RT "Hypothalamic expression of ART, a novel gene related to agouti, is
RT up-regulated in obese and diabetic mutant mice.";
RL Genes Dev. 11:593-602(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=EYE;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Plays a role in weight homeostasis. May play a role in
CC the regulation of melanocortin receptors within the hypothalamus
CC and adrenal gland, and therefore in the central control of
CC feeding.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in arcuate nucleus and median
CC eminence, adrenal gland (medulla), hypothalamus, testis, and lung.
CC -!- INDUCTION: Hypothalamic expression is elevated circa 10-fold in
CC ob/ob and db/db mice.
CC -!- SIMILARITY: Belongs to the agouti family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U89484; AAB68620.1; -; mRNA.
CC EMBL; U89486; AAB68622.1; -; Genomic_DNA.
CC EMBL; BC079902; AAT79902.1; -; mRNA.
CC HSSP; O00253; 1MR0.
CC SMR; P56473; 86-131.
CC Ensembl; ENSMUSG00000005705; Mus musculus.
CC MGI; MGI:892013; AgRP.
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0005184; F:neuropeptide hormone activity; IEA.
CC GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
CC GO; GO:0007582; P:physiological process; IEA.
CC InterPro; IPR007733; Agouti.
CC Pfam; PF05039; Agouti; 1.
KW Signal.
FT CHAIN 21 131 Agouti-related protein.
FT

```